

This Page Is Inserted by IFW Operations
and is not a part of the Official Record.

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

REVISED VERSION

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
30 May 2002 (30.05.2002)

PCT

(10) International Publication Number
WO 02/042326 A1

(51) International Patent Classification⁷: C07K 14/08,
14/115, 14/135, 16/10, A61K 38/16, 39/155, A61P 11/00,
C12N 15/45, 15/40, C12Q 1/68, G01N 33/68

(21) International Application Number: PCT/AU01/01517

(22) International Filing Date:
22 November 2001 (22.11.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/252,767 22 November 2000 (22.11.2000) US

(71) Applicant (for all designated States except US): BIOTA
SCIENTIFIC MANAGEMENT PTY LTD [AU/AU];
616 St. Kilda Road, Melbourne, Victoria 3004 (AU).

(72) Inventors; and

(75) Inventors/Applicants (for US only): MASON, Anthony,
John [AU/AU]; 509 Centre Road, Bentleigh, Victoria 3204
(AU). TUCKER, Simon, Peter [GB/AU]; 10B Karrakatta
Street, Black Rock, Victoria 3193 (AU). YOUNG, Paul,
R. [AU/AU]; 20 Joffre Street, Ashgrove, Queensland 4060
(AU).

(74) Agents: OBRANOVICH, Tania, D. et al.; Davies Colli-
son Cave, 1 Little Collins Street, Melbourne, Victoria 3000
(AU).

(81) Designated States (national): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG,
SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN,
YU, ZA, ZM, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR,
GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent
(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR,
NE, SN, TD, TG).

Published:
— with international search report

(88) Date of publication of the revised international search
report: 6 September 2002

(15) Information about Correction:
see PCT Gazette No. 36/2002 of 6 September 2002, Sec-
tion II

For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

(54) Title: A METHOD OF EXPRESSION AND AGENTS IDENTIFIED THEREBY

(57) Abstract: The present invention relates generally to a method for the *in vitro* or *in vivo* production, by a eukaryotic host cell, of a protein from a negative sense single stranded RNA virus and, more particularly, to a method for the *in vitro* or *in vivo* production by a eukaryotic host cell of a protein from a virus of the family Paramyxoviridae and agents identified thereby. Still more particularly, said protein is the F, N, P or SH protein, the encoding nucleic acid molecule of which has been optimised for expression in a eukaryotic host cell. In yet another aspect, the present invention relates to a method for modulating the functional activity of an F protein. More particularly, said modulation is predicated on modulation of the functioning of a novel intrasequence cleavage event. In still another aspect, the protein expression product produced in accordance with the optimised expression method of the present invention and the method of modulating F protein functional activity are useful in a range of applications including, but not limited to, the identification, design and/or modification of agents capable of modulating functional activity of the subject protein. The proteins, encoding nucleic acid molecules and agents identified in accordance with the present invention are useful, *inter alia*, in the treatment and/or prophylaxis of viral infections.

CORRECTED VERSION

INTERNATIONAL SEARCH REPORT

International application No.

PCT/AU01/01517

A. CLASSIFICATION OF SUBJECT MATTER		
Int. Cl. ⁷ : C07K 14/08, 14/115, 14/135, 16/10; A61K 38/16, 39/155; A61P 11/00; C12N 15/45, 15/40; C12Q 1/68; G01N 33/68;		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols)		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) PubMed, keywords: F protein, fusion, RSV, respiratory syncytial virus, Paramyxoviridae, Pneunovirus and similar terms; STN file Medline: fusion, RSV, and sequences RARR, KKRKRR; Espace, keywords, f protein, expression; ANGIS: seq id's 3, 4, 5, 6, 556, 559, 562		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	EP 780475 B1 (SCHWEIZ SERUM- & IMPFINSTITUT) 25 June 1997 see whole document	1 - 14 1 - 25, 56 - 59
X, Y	WO 99/02694 A1 (THE UNIVERSITY OF QUEENSLAND) 21 January 1999 see whole document	1- 25, 56 -59
X, Y	WO 96/09378 A1 (THE GENERAL HOSPITAL CORPORATION) 28 March 1996 see whole document	1- 25, 56 -59
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C <input checked="" type="checkbox"/> See patent family annex		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 24 January 2002		Date of mailing of the international search report 19 FEB 2002
Name and mailing address of the ISA/AU AUSTRALIAN PATENT OFFICE PO BOX 200, WODEN ACT 2606, AUSTRALIA E-mail address: pct@ipaustalia.gov.au Facsimile No. (02) 6285 3929		Authorized officer DAVID GRIFFITHS Telephone No : (02) 6283 2628

INTERNATIONAL SEARCH REPORT

International application No.

PCT/AU01/01517

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 4,619,942 A (TIDWELL, <i>et al.</i>) 28 October 1986 col. 1, line 61 - col. 2, line 16	26
X	WO 99/62932 A2 (VANDERBILT UNIVERSITY) 9 December 1999 page 12 line 15 to page 14 line 11	41, 62, 75, 79
X	SAKURAI, Hiroshi <i>et al.</i> , "Human Antibody Responses to Mature and Immature Forms of Viral Envelope in Respiratory Syncytial Virus Infection: Significance for Subunit Vaccines", <i>Journal of Virology</i> , Vol. 73, No. 4, April 1999, pp. 2956 - 2962 see whole document	41, 62, 75, 79
X	LI, Xiaomao <i>et al.</i> , "Protection against Respiratory Syncytial Virus Infection by DNA Immunization", <i>J. Exp. Med.</i> , Vol 188, No. 4, 17 August 1998, pp. 681 - 688 see whole document	41, 62, 75, 79
X	LOPEZ, Juan A. <i>et al.</i> , "Antigenic Structure of Human Respiratory Syncytial Virus Fusion Protein", <i>Journal of Virology</i> , Vol. 72, No. 8, August 1998, pp. 6922-6928 see whole document	41, 62, 75, 79
P, X	ZIMMER, G. <i>et al.</i> , "Proteolytic Activation of Respiratory Syncytial Virus Fusion Protein". <i>Journal of Biological Chemistry</i> , Vol. 276, No. 34, pp. 31642-31650 see whole document	44 - 49

INTERNATIONAL SEARCH REPORT

International application No.

PCT/AU01/01517

Box I Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos :
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☒ Claims Nos : 1 - 41 (in part), 42, 43, 44 - 79 (in part)
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
The breadth of the claims is such that a full, meaningful search is not possible on economic grounds. The search was therefore confined largely to the optimisation of expression in eukaryotes by replacing less-preferred codons by more preferred codons. A keyword search was done to cover the other claimed aspects but due to the breadth of the claims this cannot be regarded as complete. It was not technically feasible to search claims 42 and 43.
3. ☐ Claims Nos :
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a)

Box II Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

See supplemental sheet

- ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims
2. ☒ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/AU01/01517

Supplemental Box

(To be used when the space in any of Boxes I to VIII is not sufficient)

The following separate inventions have been identified:

- (i) a method of facilitating the production of viral protein (*e.g.* claim 1)
- (ii) a method of regulating the functional activity of viral F protein (*e.g.* claim 26)
- (iii) a method of detecting an agent capable of regulating the functional activity of viral F protein (*e.g.* claim 34)
- (iv) an agent capable of regulating the functional activity of viral F protein (*e.g.* claim 41)
- (v) a viral F protein variant (*e.g.* claim 44)
- (vi) a recombinant viral protein construct optimised for expression in a eukaryote (*e.g.* claim 60).

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/AU01/01517

This Annex lists the known "A" publication level patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document Cited in Search Report		Patent Family Member	
EP 780475	AU 68208/96 WO 9706270	CA 2228956	EP 846181
WO 99/02694	AU 81999/98	EP 1002091	
WO 96/09378	AU 35099/95 TR 960230 US 5795737 WO 9711086	CA 2200342 US 5786464 CA 2231394	EP 781329 ZA 9507846 EP 851868
US 4619942	US 4324794	US 4397863	
WO 99/62932	US 6315810 EP 882679	US 6299827 JP 11005069	DE 19723599
END OF ANNEX			

- 6 -

Specific mutations in amino acid sequence are represented herein as "Xaa₁nXaa₂" where Xaa₁ is the original amino acid residue before mutation, n is the residue number and Xaa₂ is the mutant amino acid. The abbreviation "Xaa" may be the three letter or single letter amino acid code. A mutation in single letter code is represented, for example, by X₁nX₂ where X₁ and X₂ are the same as Xaa₁ and Xaa₂, respectively. The amino acid residues for F protein are numbered with the first residue R in the motif RARR being residue number 106.

One aspect of the present invention is directed to a method of facilitating production of a protein or derivative thereof from a negative sense single stranded RNA virus, said method comprising expressing in a host cell a nucleic acid molecule encoding said protein or derivative thereof, the nucleotide sequence of which nucleic acid molecule is optimised for expression by a eukaryotic cell.

Another aspect of the present invention provides a method of facilitating production of a protein or derivative thereof from a virus of the family Paramyxoviridae, said method comprising expressing in a host cell a nucleic acid molecule encoding said protein or derivative thereof, the nucleotide sequence of which nucleic acid molecule is optimised for expression by a eukaryotic host cell.

Yet another aspect of the present invention provides a method of facilitating production of a protein or derivative thereof from a negative sense single stranded RNA virus, which protein directly or indirectly facilitates fusion of any one or more viral components with any one or more host cell components, said method comprising expressing in a host cell a nucleic acid molecule encoding said protein or derivative thereof, the nucleotide sequence of which nucleic acid molecule is optimised for expression by a eukaryotic cell.

Still another aspect of the present invention is therefore more particularly directed to a method of facilitating production of a F protein or derivative thereof from a negative sense single stranded RNA virus, said method comprising expressing in a host cell a nucleic acid

- 7 -

molecule encoding said F protein or derivative thereof, the nucleotide sequence of which nucleic acid molecule is optimised for expression by a eukaryotic cell.

Yet still another aspect of the present invention provides a method of facilitating
5 production of a N protein or derivative thereof from a negative sense single stranded RNA virus, said method comprising expressing in a host cell a nucleic acid molecule encoding said N protein or derivative thereof, the nucleotide sequence of which nucleic acid molecule is optimised for expression by a eukaryotic cell.

10 Still yet another aspect of the present invention provides a method of facilitating production of a P protein or derivative thereof from a negative sense single stranded RNA virus, said method comprising expressing in a host cell a nucleic acid molecule encoding said P protein or derivative thereof, the nucleotide sequence of which nucleic acid molecule is optimised for expression by a eukaryotic cell.

15 A further aspect provides a method of facilitating production of a SH protein or derivative thereof from a negative sense single stranded RNA virus, said method comprising expressing in a host cell a nucleic acid molecule encoding said SH protein or derivative thereof, the nucleotide sequence of which nucleic acid molecule is optimised for
20 expression by a eukaryotic cell.

Another further aspect provides a method of facilitating production of a protein or derivative thereof from a negative sense single stranded RNA virus, said method comprising expressing in a mammalian host cell a nucleic acid molecule encoding said
25 protein or derivative thereof, the nucleotide sequence of which nucleic acid molecule is optimised for expression by said mammalian host cell.

Yet another further aspect of the present invention is directed to a method of facilitating production of a protein or derivative thereof from a negative sense single stranded RNA
30 virus, said method comprising expressing in a mammalian host cell a nucleic acid molecule encoding said protein or derivative thereof, the nucleotide sequence of which

- 8 -

nucleic acid molecule is optimised for expression by said mammalian host cell wherein said optimisation is codon optimisation and/or nucleotide splice site deletion.

5 Still another further aspect provides a method of facilitating production of F protein or derivative thereof from respiratory syncytial virus, said method comprising expressing in a mammalian host cell a nucleic acid molecule encoding said F protein or derivative thereof, the nucleotide sequence of which nucleic acid molecule is optimised for expression by said mammalian host cell wherein said optimisation is nucleotide splice site deletion.

10 Still yet another further aspect of the present invention is directed to a method of facilitating production of a F_{sol} portion of an F protein or derivative thereof from respiratory syncytial virus, said method comprising expressing in a host cell a nucleic acid molecule encoding said F_{sol} portion or derivative thereof, the nucleotide sequence of which nucleic acid molecule is optimised for expression by said mammalian host cell wherein
15 said optimisation is nucleotide splice site deletion.

Yet still another further aspect provides a method of facilitating production of F protein or derivative thereof from respiratory syncytial virus, said method comprising expressing in a mammalian host cell a nucleic acid molecule encoding said F protein or derivative thereof,
20 the nucleotide sequence of which nucleic acid molecule is optimised for expression by said mammalian host cell wherein said optimisation is codon optimisation.

Another aspect of the present invention is directed to a method of facilitating production of a F_{sol} portion of an F protein or derivative thereof from respiratory syncytial virus, said
25 method comprising expressing in a host cell a nucleic acid molecule encoding said F_{sol} portion or derivative thereof, the nucleotide sequence of which nucleic acid molecule is optimised for expression by said mammalian host cell wherein said optimisation is codon optimisation.

30 Yet another aspect of the present invention provides a method of facilitating production of F protein or derivative thereof from respiratory syncytial virus, said method comprising

- 9 -

expressing in a mammalian host cell a nucleic acid molecule encoding said F protein or derivative thereof, the nucleotide sequence of which nucleic acid molecule is optimised for expression by said mammalian host cell wherein said optimisation is nucleotide splice site deletion and codon optimisation.

5

Still another another aspect of the present invention provides a method of facilitating the production of a F protein or derivative thereof from a respiratory syncytial virus, said method comprising expressing in a host cell the nucleotide sequence set forth in <400>5 or derivative thereof.

10

Yet still another aspect provides a method of facilitating the production of a F_{sol} portion of an F protein or derivative thereof from respiratory syncytial virus, said method comprising expressing in a host cell the nucleotide sequence set forth in <400>6 or derivative thereof.

15

Still yet another aspect provides a method of facilitating production of P protein or derivative thereof from respiratory syncytial virus, said method comprising expressing in a mammalian host cell a nucleic acid molecule encoding said P protein or derivative thereof, the nucleotide sequence of which nucleic acid molecule is optimised for expression by said mammalian host cell wherein said optimisation is codon optimisation.

20

A further aspect provides a method of facilitating the production of a P protein or derivative thereof from respiratory syncytial virus, said method comprising expressing in a host cell the nucleotide sequence set forth in <400>556 or derivative thereof.

25

Another further aspect provides a method of facilitating production of N protein or derivative thereof from respiratory syncytial virus, said method comprising expressing in a mammalian host cell a nucleic acid molecule encoding said N protein or derivative thereof, the nucleotide sequence of which nucleic acid molecule is optimised for expression by said mammalian host cell wherein said optimisation is codon optimisation.

30

- 10 -

Yet another further aspect provides a method of facilitating the production of a N protein or derivative thereof from respiratory syncytial virus, said method comprising expressing in a host cell the nucleotide sequence set forth in <400>559 or derivative thereof.

- 5 Still another further aspect provides a method of facilitating production of SH protein or derivative thereof from respiratory syncytial virus, said method comprising expressing in a mammalian host cell a nucleic acid molecule encoding said SH protein or derivative thereof, the nucleotide sequence of which nucleic acid molecule is optimised for expression by said mammalian host cell wherein said optimisation is codon optimisation.

10

Still yet another further aspect provides a method of facilitating the production of a SH protein or derivative thereof from respiratory syncytial virus, said method comprising expressing in a host cell the nucleotide sequence set forth in <400>562 or derivative thereof.

15

In another aspect, the present invention should be understood to extend to the optimised nucleic acid molecules described herein and to the expression products derived therefrom.

- Yet another aspect of the present invention is directed to a method of regulating the functional activity of a viral F protein, which protein in its non-fully functional form comprises an F2 portion linked, bound or otherwise associated with an F1 portion, which F2 portion comprises an intervening peptide sequence, said method comprising modulating cleavage of said intervening peptide sequence wherein excision of at least part of said intervening sequence from said non-fully functional form of said F protein up-regulates F protein functional activity.
- 20
- 25

- Still another aspect of the present invention is directed to a method of regulating the functional activity of a Paramyxoviridae derived F protein, which protein in its non-fully functional form comprises an F2 portion linked, bound or otherwise associated with an F1 portion, which F2 portion comprises an intervening peptide sequence, said method comprising modulating cleavage of said intervening peptide sequence wherein excision of
- 30

- 11 -

at least part of said intervening sequence from said non-fully functional form of said F protein up-regulates F protein functional activity.

Yet still another aspect of the present invention provides a method of regulating the functional activity of a respiratory syncytial virus F protein, which protein in its non-fully functional form comprises an F2 portion linked, bound or otherwise associated with an F1 portion, which F2 portion comprises an intervening peptide sequence, said method comprising modulating cleavage of said intervening peptide sequence, wherein excision of at least part of said intervening sequence from said non-fully functional form of said F protein up-regulates F protein functional activity and wherein said cleavage events occur at the cleavage sites defined by the peptide sequences RARR (<400>564) and KKRKRR (<400>563).

In a related aspect, the present invention provides a method of regulating the functional activity of a viral F protein, which protein in its non-fully functional form comprises the structure:

X_1, X_2, X_3

wherein:

X_1 comprises the non-intervening peptide sequence region of the F2 portion;

X_2 comprises the intervening peptide sequence region of the F₂ portion; and

X_3 comprises the F1 portion

said method comprising modulating cleavage of said intervening peptide sequence wherein excision of at least part of said intervening sequence from said non-fully functional form of said F protein up-regulates F protein functional activity.

Still yet another aspect provides a method of inhibiting, retarding or otherwise down-regulating the functional activity of a Paramyxoviridae derived F protein, which protein in its non-fully functional form comprises an F2 portion linked, bound or otherwise

- 12 -

associated with an F1 portion, which F2 portion comprises an intervening peptide sequence, said method comprising inhibiting or otherwise down-regulating cleavage of said intervening peptide sequence.

- 5 A further aspect of the present invention provides a method of down-regulating the functional activity of a Paramyxoviridae derived F protein, which protein in its non-fully functional form comprises the structure:

$X_1X_2X_3$

10

wherein:

X_1 comprises the non-intervening peptide sequence region of the F2 portion;

X_2 comprises the intervening peptide sequence region of the F2 portion; and

X_3 comprises the F1 portion

15

said method comprising inhibiting or otherwise down-regulating cleavage of said intervening peptide sequence.

- Another further aspect provides a method for detecting an agent capable of regulating the functional activity of a viral F protein or derivative thereof said method comprising
20 contacting a eukaryotic cell expressing an optimised nucleic acid molecule encoding said viral F protein or derivative thereof, as hereinbefore described, with a putative modulatory agent and detecting an altered expression phenotype and/or functional activity.

- 25 In yet another aspect there is provided a method for detecting an agent capable of regulating the functional activity of a viral F protein or derivative thereof said method comprising contacting a host cell, which host cell expresses a nucleic acid molecule encoding the non-fully functional form of said viral F protein or derivate thereof as hereinbefore described, with a putative modulatory agent and detecting an altered
30 expression phenotype and/or altered functional activity wherein said agent modulates cleavage of the intervening peptide sequence.

- 13 -

Still another further aspect of the present invention is directed to a method for analysing, designing and/or modifying an agent capable of interacting with a viral F protein or derivative thereof and modulating at least one functional activity associated with said
5 protein, which protein is produced in accordance with the method of the present invention said method comprising contacting said F protein or derivative thereof with a putative agent and assessing the degree of interactive complementarity of said agent with said protein.

Still yet another further aspect of the present invention is directed to an agent capable of
10 interacting with a viral F protein and modulating at least one functional activity associated with said viral protein.

In still another aspect there is provided a viral F protein variant comprising a mutation in the intervening peptide sequence wherein said variant exhibits modulated functional
15 activity relative to wild type F protein or a derivative, homologue, analogue, chemical equivalent or mimetic of said variant.

Another aspect of the present invention provides a viral F protein variant comprising a mutation in the intervening peptide sequence wherein said variant exhibits down-regulated
20 functional activity relative to wild type F protein or a derivative, homologue, analogue, chemical equivalent or mimetic of said variant.

Yet another aspect provides a respiratory syncytial virus F protein variant comprising a mutation in the cleavage site defined by amino acids RARR (<400>564) wherein said
25 variant exhibits down-regulated functional activity relative to wild type F protein or a derivative, homologue, analogue, chemical equivalent or mimetic of said variant.

Preferably said mutation comprises one or more of the amino acid substitutions selected from the following list:

30

(i) R106G

- 14 -

(ii) A107Q

(iii) R108G

Still more preferably said F protein variant comprises the sequence substantially as set
5 forth in <400>565.

Still another aspect provides a respiratory syncytial virus F protein variant comprising a
multiple amino acid deletion from the intervening peptide sequence wherein said variant
exhibits down-regulated functional activity relative to wild type F protein or a derivative,
10 homologue, analogue, chemical equivalent of said variant.

It is more preferably provided that said amino acid deletion is a partial deletion of the
intervening peptide sequence and more preferably a deletion of the peptide sequence

15 RARRELPRFMNYTLNNAKKTNVTLS <400>569.

Still more preferably said variant comprises the amino acid sequence substantially as set
forth in <400>567.

20 Yet still another aspect of the present invention is directed to an isolated nucleic acid
molecule selected from the list consisting of:

(i) An isolated nucleic acid molecule or derivative or equivalent thereof comprising a
nucleotide sequence encoding or complementary to a sequence encoding a viral F
protein variant or derivative, homologue, analogue, chemical equivalent or mimetic
25 of said variant, which variant comprises a mutation in the intervening peptide
sequence wherein said variant exhibits modulated functional activity relative to
wild-type F protein.

(ii) An isolated nucleic acid molecule or derivative or equivalent thereof comprising a
30 nucleotide sequence encoding or complementary to a sequence encoding a viral F
protein variant or derivative, homologue, analogue, chemical equivalent or mimetic

- 15 -

of said variant, which variant comprises a mutation in the intervening peptide sequence wherein said variant exhibits down-regulated functional activity relative to wild-type F protein.

- 5 (iii) An isolated nucleic acid molecule or derivative or equivalent thereof comprising a nucleotide sequence encoding or complementary to a sequence encoding a respiratory syncytial virus F protein or derivative, homologue, analogue, chemical equivalent or mimetic of said variant, which variant comprises a mutation in the cleavage site defined by amino acids RARR wherein said variant exhibits down-regulated functional activity relative to wild-type F protein.
- 10
- (iv) An isolated nucleic acid molecule or derivative or equivalent thereof comprising a nucleotide sequence encoding or complementary to a sequence encoding a respiratory syncytial virus F protein variant or derivative, homologue, analogue, chemical equivalent or mimetic of said variant, which variant comprises one or more of the amino acid substitutions selected from the following list:
- 15
- (a) R106G
 - (b) A107Q
 - (c) R108G
- 20
- (v) An isolated nucleic acid molecule or derivative or analogue thereof comprising a nucleotide sequence encoding or complementary to a sequence encoding a viral F protein variant or derivative, homologue, analogue, chemical equivalent or mimetic of said variant, which variant comprises a multiple amino acid deletion from the intervening peptide sequence wherein said variant exhibits down-regulated functional activity relative to wild-type F protein.
- 25
- (vi) An isolated nucleic acid molecule or derivative or analogue thereof comprising a nucleotide sequence encoding or complementary to a sequence encoding a viral F protein variant or derivative, homologue, analogue, chemical equivalent or mimetic
- 30

- 16 -

of said variant, which variant comprises a partial deletion of the intervening peptide sequence and more preferably a deletion of the peptide sequence

RARRELPRFMNYTLNNAKKTNTLS <400>569.

5

(vii) An isolated nucleic acid molecule or derivative or analogue thereof comprising a nucleotide sequence encoding or complementary to a sequence encoding a viral F protein variant or derivative, homologue, analogue, chemical equivalent or mimetic of said variant, which variant comprises the amino acid sequence substantially as set forth in <400>567.

10

(viii) An isolated nucleic acid molecule or derivative or analogue thereof comprising a nucleotide sequence encoding or complementary to a sequence encoding a viral F protein variant or derivative, homologue, analogue, chemical equivalent or mimetic of said variant, which variant comprises the amino acid sequence substantially as set forth in <400>565.

15

(ix) An isolated nucleic acid molecule or derivative or analogue thereof comprising the nucleotide substantially as set forth in <400>568.

20

(x) An isolated nucleic acid molecule or derivative or analogue thereof comprising the nucleotide substantially as set forth in <400>566.

Still yet another aspect of the present invention provides a recombinant viral construct comprising a nucleic acid molecule encoding a viral F protein or derivative thereof, the nucleotide sequence of which nucleic acid molecule comprises codons optimised for expression in a eukaryotic cell, wherein said recombinant viral construct is effective in inducing, enhancing or otherwise stimulating an immune response to said F protein.

A further aspect of the present invention provides a recombinant viral construct comprising a nucleic acid molecule encoding a viral F protein variant or derivative thereof wherein

30

- 17 -

said recombinant viral construct is effective in inducing, enhancing or otherwise stimulating an immune response to said F protein variant.

Another further aspect of the present invention relates to a vaccine comprising a
5 recombinant viral construct which construct comprises a nucleic acid molecule encoding a
respiratory syncytial virus F protein or derivative thereof, the nucleotide sequence of which
nucleic acid molecule is optimised for expression in a eukaryotic cell wherein said
recombinant viral construct is effective in inducing, enhancing or otherwise stimulating an
immune response to said F protein.

10

Yet another further aspect of the present invention relates to a vaccine comprising a
recombinant viral construct which construct comprises a nucleic acid molecule encoding a
respiratory syncytial virus F protein variant or derivative thereof, wherein said
recombinant viral construct is effective in inducing, enhancing or otherwise stimulating an
15 immune response to said F protein variant.

In accordance with these aspects of the present invention, the nucleotide sequence of the
subject nucleic acid molecule is preferably the nucleotide sequence defined in <400>5,
<400>6, <400>566 or <400>568.

20

Still another further aspect of the present invention provides the method of modulating at
least one functional activity associated with a viral F protein in a subject, said method
comprising introducing into said subject and effective amount of an F protein modulatory
agent for a time and under condition sufficient for said agent to interact with said F protein.

25

Still yet another further aspect of the present invention provides a method of modulating at
least one functional activity associated with a viral F protein, said method comprising
contacting said viral F protein with an effective amount of an F protein modulatory agent
for a time and under conditions sufficient for said agent to interact with said F protein.

30

- 18 -

Yet still another further aspect of the present invention relates to a method for the treatment and/or prophylaxis of a condition characterised by infection with a negative sense single stranded RNA virus in a subject, said method comprising administering to said subject an effective amount of an agent, which agent is capable of down-regulating at least one functional activity of the F protein expressed by said virus, for a time and under conditions sufficient for said agent to interact with said F protein.

In still yet another aspect, the present invention relates to a method for the treatment and/or prophylaxis of a condition characterised by infection with a negative sense single stranded RNA virus in a subject, said method comprising administering to said subject an effective amount of a composition comprising an F protein or derivative thereof, F protein variant or derivative thereof and/or a nucleic acid molecule encoding said F protein or F protein variant as hereinbefore defined or a derivative, homologue, analogue, chemical equivalent or mimetic of said protein or nucleic acid molecule for a time and under conditions sufficient for said composition to down-regulate said viral F protein functional activity.

In another aspect the present invention relates to the use of an agent capable of modulating at least one functional activity of a viral F protein, which agent is identified and/or generated in accordance with the methods hereinbefore defined, in the manufacture of a medicament for the treatment and/or prophylaxis of a condition characterised by infection with a negative sense single stranded RNA virus.

In still another aspect the present invention relates to the use of a composition comprising an F protein or derivative thereof, F protein variant or derivative thereof, nucleic acid molecule encoding said F protein or F protein variant as hereinbefore defined or a derivative, homologue, analogue, chemical equivalent or mimetic of said protein or nucleic acid molecule, in the manufacture of a medicament for the treatment and/or prophylaxis of a condition characterised by infection with a negative sense single stranded RNA virus.

- 19 -

In another aspect the present invention relates to the use of an agent, which agent is identified in accordance with the methods hereinbefore defined, in the manufacture of a medicament for the modulation of at least one viral F protein associated functional activity.

- 5 Yet another aspect relates to agents for use in modulating the functional activity of a viral F protein wherein said agent is identified in accordance with the methods hereinbefore defined.

- Still yet another aspect relates to agents for use in the treatment and/or prophylaxis of a
10 condition characterised by infection with a negative sense single stranded RNA virus wherein said agent is identified in accordance with the methods hereinbefore defined.

- Yet still another aspect relates to a composition comprising an F protein or derivative thereof, F protein variant or derivative thereof, a nucleic acid molecule encoding said F
15 protein or F protein variant as hereinbefore defined or a derivative, homologue, analogue, chemical equivalent or mimetic of said protein or nucleic acid molecule for use in the treatment and/or prophylaxis of a condition characterised by infection with a negative sense single stranded RNA virus.

- 20 In yet another aspect the present invention relates to a pharmaceutical composition comprising an active ingredient, as hereinbefore defined, and one or more pharmaceutically acceptable carriers and/or diluents.

25

- 20 -

Single and three letter abbreviations used throughout the specification are defined in Table 2.

TABLE 2**Single and three letter amino acid abbreviations**

5

Amino Acid		Three-letter Abbreviation	One-letter Symbol
10	Alanine	Ala	A
	Arginine	Arg	R
	Asparagine	Asn	N
	Aspartic acid	Asp	D
	Cysteine	Cys	C
15	Glutamine	Gln	Q
	Glutamic acid	Glu	E
	Glycine	Gly	G
	Histidine	His	H
	Isoleucine	Ile	I
20	Leucine	Leu	L
	Lysine	Lys	K
	Methionine	Met	M
	Phenylalanine	Phe	F
	Proline	Pro	P
25	Serine	Ser	S
	Threonine	The	T
	Tryptophan	Trp	W
	Tyrosine	Tyr	Y
	Valine	Val	V
30	Any residue	Xaa	X

- 21 -

BRIEF DESCRIPTION OF THE DRAWINGS

- Figure 1a** is a schematic representation of the 574 amino acid sequence of the human RSV fusion protein F. Amino acid numbers 1-22 comprises the signal sequence. The F2 subunit comprises amino acid numbers 23-130. The fusion cleavage (site 1) is amino acid numbers 131-136. Site 2 comprises residues 106-109. The F1 subunit comprises residues 136-574. The transmembrane domain is believed to span residues 525-548. The cytoplasmic domain comprises residues 549-574.
- Figure 1b** is a schematic representation of the amino acid sequence of the 524 residue soluble F protein. This protein is referred to as F_{sol} . F_{sol} is formed by expressing the coding sequence for F minus the residues encoding the transmembrane domain and the cytoplasmic domain of F.
- Figure 1c** is a schematic representation of F and F_{sol} . Cleavage positions of site 1 and site 2 are designated. Hydrophobic regions are shaded in black (from left to right, signal sequence, fusion peptide and transmembrane domain). Downward facing flags designate positions of potential N-linked glycosylation sites. The 24 amino acid region bounded by cleavage sites 1 and 2 is shown as a cross-hatched region.
- Figure 2a** is a schematic representation of the alignment of sequences coding for the human RSV F protein. F.viral refers to the sequence as found in wild type A2 RSV strain. F refers to the sequence which differs in 27/1725 positions from the viral sequence. Those changes were made in order to introduce unique restriction sites to the sequence. F.opt. refers to the F coding sequence which has been changed to allow for higher expression levels as outlined in the accompanying application. A total of 378/1725 nucleotides have been changed from the F.viral sequence. Underneath the boxed sequence a consensus sequence is shown.

- 22 -

Figure 2b is a schematic representation of the alignment of sequences coding for the human RSV F_{sol} protein. F_{sol}.viral refers to the sequence as found in the wild type A2 RSV strain. F_{sol} refers to the sequence which differs from the viral sequence in 24/1575 nucleotides. All of these changes were incorporated to introduce unique restriction sites.

5 F_{sol}.opt. refers to the F_{sol} coding sequence optimised as described herein. A total of 334/1575 nucleotides have been changed. A consensus is shown under the boxed sequences.

10 **Figures 3a and b** are schematic representations of the DNA sequences optimised for expression as cloned in the expression vector pCICO.F.FL.opt (a) and pCICO.F.opt (b). The plasmid pCICO.F.FL.opt contains the sequence referred to in Figure 2a as F.opt.. The plasmid pCICO.F.opt contains the sequence referred to in Figure 2b as F_{sol}.opt. 5' and 3' untranslated sequences not included in the Figure 2 sequences are shown in this Figure.

15 **Figures 4a and b** are schematic representations of the construction of F and F_{sol} expression vectors. These diagrams describe in detail the steps involved in constructing expression vectors pCICO.F.FL.opt and pCICO.F.opt. See text of examples for details. As previously noted pCICO.F.FL.opt contains the optimised sequence F.opt. (Figure 2a) and pCICO.F.opt contains the optimised sequence F_{sol}.opt (Figure 2b).

20

Figure 5 is an image of an autoradiograph of a 10% SDS-PAGE gel of a immunoprecipitation of 35-S labelled supernatants from 293 cells transfected with lane (a) pCICO.FS3 (containing viral F_{sol} sequence) lane (b) pCICO.F.opt (containing optimised F_{sol} sequence). Lane (c) is from mock-transfected cells. Lane (d) contains radioactively
25 labelled molecular weight markers. The F_{sol} protein migrates at approximately 60 kd in size.

Figure 6 is a schematic representation of the alignment of sequences coding for the human RSV F protein. F.viral refers to the sequence as found in wild type A2 RS strain
30 (<400>571). F.nat refers to the sequence found in a RSV A2 cDNA clone assembled in these studies (<400>572). The two sequences differ in two places (nt 174 and 222) which

- 23 -

does not effect the coding potential. Underneath the boxed sequence a consensus sequence is shown (<400>573).

5 **Figure 7** is a western blot of protein samples derived from 293 cells transfected with WT (pCICO.F.FL.opt), A2 (pCICO.F.nat) and Ctrl (control) plasmids. Cells were havested at 24, 48 and 72 hours post transfection. Cell lysates were analysed by 12% polyacrylamide SDS-PAGE and after electrophoresis proteins were electroblotted onto a nitrocellulose membrane. F protein was detected as described in example 5. The immuno-reactive F bands F1 and F1' are indicated by arrows. The position of molecular weight markers is
10 shown.

Figure 8 is photographs of 293 cells transfected with pCICO.F.FL.opt (a), pCICO.F.nat (b) and control plasmid (c). Photographs were taken 48 hours post transfection and the magnification is 400X. Figures a, b and c flow from top to bottom.

- 24 -

DETAILED DESCRIPTION OF THE INVENTION

The present invention is predicated, in part, on the development of a negative sense single stranded RNA viral protein expression system based on optimisation of expression of the viral protein encoding nucleic acid sequence such that expression of the subject nucleic acid molecule sequence by a given eukaryotic host cell is facilitated and/or improved. In a related aspect, the inventors have identified a novel cleavage site in the F viral protein, the cleavage of which is thought to be essential for the generation of a fully functionally active F protein. These developments now permit the recombinant production of viral proteins and the identification and design of agents for use in modulating functional activity of the subject proteins.

Accordingly, one aspect of the present invention is directed to a method of facilitating production of a protein or derivative thereof from a negative sense single stranded RNA virus, said method comprising expressing in a host cell a nucleic acid molecule encoding said protein or derivative thereof, the nucleotide sequence of which nucleic acid molecule is optimised for expression by a eukaryotic cell.

Reference to a "negative sense single stranded RNA virus" should be understood as a reference to any negative sense single stranded RNA virus, and includes, but is not limited to, viruses of the family Paramyxoviridae, Rhabdoviridae, Filoviridae, Orthomyxoviridae, Bunyaviridae or Arenaviridae. Preferably, said negative sense single stranded RNA virus is of the family Paramyxoviridae. Without limiting the present invention to any one theory or mode of action, viruses of the family Paramyxoviridae are cytoplasm replicating viruses. In this regard, RNA replication involves mRNA transcription from the genomic RNA via the virion transcriptase. Utilising the protein products of this transcription, there follows the production of a full length positive stranded template which is used for the synthesis of genomic RNA. The genome is transcribed from the the 3' end by virion associated enzymes into mRNAs. Replication takes place in the cytoplasm and assembly occurs via budding on the plasma membrane. The subject budding occurs through the host cell plasma membrane at sites containing the virus envelope proteins.

- 25 -

Accordingly, there is more particularly provided a method of facilitating production of a protein or derivative thereof from a virus of the family Paramyxoviridae, said method comprising expressing in a host cell a nucleic acid molecule encoding said protein or derivative thereof, the nucleotide sequence of which nucleic acid molecule is optimised for
5 expression by a eukaryotic host cell.

Still more preferably, said virus is of the sub-family Pneumovirinae and most preferably said virus is respiratory syncytial virus.

10 Reference to a "protein from a negative sense single stranded RNA virus" should be understood as a reference to any protein which is expressed by the subject virus or a derivative of said protein. Examples of proteins include, but are not limited to, nucleocapsid associated proteins such as RNA binding proteins (e.g. N, NP), phosphoproteins (e.g. P), polymerase proteins (e.g. L), or envelope proteins (e.g. F, G, H,
15 HN or SH). It should be understood that the subject protein may exist, in its naturally occurring form, either in isolation or fused or otherwise linked to any other proteinaceous or non-proteinaceous molecule. Preferably, the subject protein is a fusion protein, N, P or SH.

20 Accordingly, in one embodiment there is provided a method of facilitating production of a protein or derivative thereof from a negative sense single stranded RNA virus, which protein directly or indirectly facilitates fusion of any one or more viral components with any one or more host cell components, said method comprising expressing in a host cell a nucleic acid molecule encoding said protein or derivative thereof, the nucleotide sequence
25 of which nucleic acid molecule is optimised for expression by a eukaryotic cell.

Reference to a viral protein which "directly or indirectly facilitates fusion of any one or more viral components with any one or more host cell components" should be understood as a reference to any viral protein which functions to induce or otherwise contribute to the
30 fusion of one or more viral molecules (such as a protein or structural component) with any one or more host cell molecules. It should be understood that this activity may comprise

- 26 -

any one of a number of functional activities attributable to the subject protein, which other activities are not necessarily related to fusion. It should also be understood that the subject functional activity may either directly facilitate fusion or it may induce or otherwise contribute to the functioning of an unrelated molecule, which unrelated molecule directly
5 facilitates the subject fusion. Preferably the viral protein is an F protein.

This embodiment of the present invention is therefore more particularly directed to a method of facilitating production of a F protein or derivative thereof from a negative sense single stranded RNA virus, said method comprising expressing in a host cell a nucleic acid
10 molecule encoding said F protein or derivative thereof, the nucleotide sequence of which nucleic acid molecule is optimised for expression by a eukaryotic cell.

Reference to a "F protein" should be understood as a reference to the viral molecule which, *inter alia*, facilitates fusion between the virus envelope and the host cell plasma membrane
15 of infected cells. The term "F protein" should be understood to encompass all forms of F protein including, for example, any mutant, polymorphic or homologous forms of F protein. Without limiting the present invention in any way, the F protein generally comprises, at the amino terminus, an F2 portion which is linked to an F1 portion. The F1 contains a transmembrane region of the molecule which is, in turn, linked to an
20 extracellular portion of the F protein. The cytoplasmic portion of the F protein comprises the carboxy terminus. As detailed earlier, the F protein is generally synthesised in a precursor form which is activated by proteolytic cleavage at the F2/F1 junction. It is though that this cleavage step reveals a fusion peptide which interacts with the host cell. The F2/F1 junction of the respiratory syncytial virus F protein is shown in Figure 1.

25

In another embodiment there is provided a method of facilitating production of a N protein or derivative thereof from a negative sense single stranded RNA virus, said method comprising expressing in a host cell a nucleic acid molecule encoding said N protein or derivative thereof, the nucleotide sequence of which nucleic acid molecule is optimised for
30 expression by a eukaryotic cell.

- 27 -

In yet another preferred embodiment there is provided a method of facilitating production of a P protein or derivative thereof from a negative sense single stranded RNA virus, said method comprising expressing in a host cell a nucleic acid molecule encoding said P protein or derivative thereof, the nucleotide sequence of which nucleic acid molecule is
5 optimised for expression by a eukaryotic cell.

In still yet another preferred embodiment there is provided a method of facilitating production of a SH protein or derivative thereof from a negative sense single stranded RNA virus, said method comprising expressing in a host cell a nucleic acid molecule
10 encoding said SH protein or derivative thereof, the nucleotide sequence of which nucleic acid molecule is optimised for expression by a eukaryotic cell.

Preferably, the negative sense single stranded RNA virus of these preferred embodiments of the present invention is a virus of the family Paramyxoviridae. More preferably the
15 virus is of the sub-family Pneumovirinae and still more preferably the subject virus is a virus of the genus Pneumovirus. Most preferably, the virus is respiratory syncytial virus.

To the extent that it is not otherwise specified, reference to a viral "protein" extends to derivatives thereof.

20

"Derivatives" of the subject protein include fragments, parts, portions, mutants, variants and mimetics thereof including fusion proteins. Parts or fragments include, for example, active regions of the subject protein. In one aspect of the present invention, for example, the subject protein is a F protein which does not comprise the transmembrane and
25 cytoplasmic portions (herein referred to as F_{sol}). The F_{sol} fragment of the F protein is useful for X-ray crystallography and other forms of modelling for purposes such as rational drug design. Derivatives may be derived from insertion, deletion or substitution of amino acids. Amino acid insertional derivatives include amino and/or carboxylic terminal fusions as well as intrasequence insertions of single or multiple amino acids. Insertional amino
30 acid sequence variants are those in which one or more amino acid residues are introduced into a predetermined site in the protein although random insertion is also possible with

- 28 -

suitable screening of the resulting product. Deletional variants are characterised by the removal of one or more amino acids from the sequence. Substitutional amino acid variants are those in which at least one residue in the sequence has been removed and a different residue inserted in its place. An example of substitutional amino acid variants are
5 conservative amino acid substitutions. Conservative amino acid substitutions typically include substitutions within the following groups: glycine and alanine; valine, isoleucine and leucine; aspartic acid and glutamic acid; asparagine and glutamine; serine and threonine; lysine and arginine; and phenylalanine and tyrosine. Additions to amino acid sequences include fusions with other peptides, polypeptides or proteins.

10

The derivatives include fragments having particular portions of the entire protein fused to peptides, polypeptides or other proteinaceous or non-proteinaceous molecules.

"Mutants" include variants of the subject protein which variants exhibit modified
15 sequences, structures and/or functions. For example, the F protein variants described herein, which variants exhibit amino acid sequence alterations leading to altered cleavage properties, fall within the scope of the term "mutants".

The term "protein" should be understood to encompass peptides, polypeptides and
20 proteins. The protein may be glycosylated or unglycosylated and/or may contain a range of other molecules fused, linked, bound or otherwise associated to the protein such as amino acids, lipids, carbohydrates or other peptides, polypeptides or proteins. Reference hereinafter to a "protein" includes a protein comprising a sequence of amino acids as well as a protein associated with other molecules such as amino acids, lipids, carbohydrates or
25 other peptides, polypeptides or proteins.

The method of the present invention is predicated on the production of a viral protein by expressing a nucleic acid molecule as herein described. In this regard, the term
"expressing" should be understood to refer to the transcription and translation of a nucleic
30 acid molecule resulting in the synthesis of a peptide, polypeptide or protein expression

- 29 -

product. The synthesis of an expression product via the translation step of nucleic acid molecule expression is herein referred to as "production" of that expression product.

The viral protein encoding nucleic acid molecule of the present invention is expressed in a
5 eukaryotic host cell. By "host cell" is meant any eukaryotic cell which can be transformed or transfected with a nucleotide sequence. Preferred eukaryotic host cells are mammalian cells and even more preferably 293 cells and Chinese Hamster Ovary cells.

Accordingly, there is provided a method of facilitating production of a protein or
10 derivative thereof from a negative sense single stranded RNA virus, said method comprising expressing in a mammalian host cell a nucleic acid molecule encoding said protein or derivative thereof, the nucleotide sequence of which nucleic acid molecule is optimised for expression by said mammalian host cell.

15 Preferably, the subject protein is a fusion protein (more particularly the F protein), N, P or SH.

Preferably, the negative sense single stranded RNA virus of these preferred embodiments of the present invention is a virus of the family Paramyxoviridae. More preferably the
20 virus is of the sub-family Pneumovirinae and still more preferably the subject virus is a virus of the genus Pneumovirus. Most preferably, the virus is respiratory syncytial virus.

The nucleic acid molecule which is expressed in accordance with the method of the present invention encodes a viral protein or derivative thereof. By "encodes" is meant that the
25 expression product comprises the subject protein or derivative. However, it should be understood that this is not intended as a restriction in any way on the diversity of the subject expression product other than that it should comprise the subject protein or derivative thereof. For example, the nucleic acid molecule which is introduced into the host cell may encode the protein fused to another protein, peptide or polypeptide (which is
30 consistent with the definition of protein "derivative" as hereinbefore provided) or the

- 30 -

nucleic acid molecule may encode multiple proteins wherein at least one of those proteins is the subject protein or derivative thereof.

Reference to the subject nucleic acid molecule being "optimised" for expression by a eukaryotic host cell should be understood as a reference to a nucleic acid molecule which has been mutated or otherwise varied such that its recombinant expression by a eukaryotic host cell is facilitated. Said "facilitation" includes, but is not limited to, inducing or improving levels of protein expression and/or functional activity of the expression product. Preferably, said optimisation takes the form of codon optimisation and/or nucleotide splice site deletion.

By "codon optimisation" is meant that at least one codon of the naturally occurring viral protein encoding nucleotide sequence has been altered such that it encodes the same amino acid as the naturally occurring codon but uses an alternative codon to that which naturally encodes the subject amino acid, which alternative codon form is more preferably expressed by a eukaryotic cell than the naturally occurring codon form.

The present invention is exemplified herein with respect to the F, P, N and SH proteins, the naturally occurring encoding nucleic acid sequences of which are defined in <400>1, <400>505, <400>508 and <400>511, respectively. Without limiting the present invention to any one theory or mode of action, the inventors have determined that eukaryotic expression of a viral gene becomes possible where selected A rich and T rich regions of the naturally occurring gene are modified to express increased numbers of G rich and C rich nucleotides. This is achieved by replacing selected A or T nucleotides with a G or C nucleotide, respectively. The resultant modified codon, however, preferably encodes the same amino acid as that encoded by the naturally occurring codon. With respect to the F gene, for example, the codon TTG commences at nucleotide 7 of the naturally occurring respiratory syncytial viral F protein encoding nucleic acid sequence (provided in <400>1). This codon encodes an L amino acid. In the codon optimised F protein encoding nucleic acid sequence, represented herein in Figure 2a, the codon TTG is modified to read CTG, which modified codon nevertheless encodes the L amino acid. The present invention does

- 31 -

not, however, relate to the sequence as published by Kuhnle *et al* (1998) insofar as the sequence is used for codon optimisation.

The preferred embodiment of the present invention is to optimise the viral protein
5 encoding nucleotide sequence such that the naturally occurring viral protein amino acid
sequence or fragment thereof, is produced. However, it should be understood that it is
nevertheless within the scope of the present invention to optimise a viral protein encoding
nucleotide sequence in terms of expressing increased G plus C content, as required to
achieve efficient mammalian host cell expression, despite the fact that an optimised codon
10 may thereafter encode an amino acid different to that originally encoded by the codon
which naturally existed at that position. This may occur, for example, where the newly
substituted amino acid does not significantly alter the structural and/or functional
properties which are required of the recombinantly produced protein. For example, certain
conservative amino acid substitutions may not alter functional properties. Similarly, amino
15 acid substitutions in regions outside the protein's functionally active regions may be
acceptable in terms of the use to which the expressed protein is to be put.

In terms of optimising the naturally occurring F protein encoding nucleotide sequence, the
number of codons which are optimised in any given situation will depend on the object to
20 be achieved. For example, optimisation of between 1 and 10 codons may be sufficient to
enable production of a level of eukaryotic host cell expression sufficient for a particular
purpose. However, in order to achieve still more efficient levels of expression and/or
expression product functional activity, it may be desirable to optimise a larger number of
codons. In this regard, in a most preferred embodiment, the optimised F, P, N and SH
25 protein encoding nucleic acid sequences correspond to the sequences defined in <400>5,
<400>556, <400>559, and <400>562, respectively. However, it should be understood that
the present invention extends to the use of derivatives of these sequences.

By "nucleotide splice site deletion" optimisation is meant that the nucleotide sequence
30 encoding a subject viral protein has been altered to remove one or more potential RNA
splice sites. Without limiting the present invention to any one theory or mode of action, it

- 32 -

- is thought that inefficient expression of nucleotide sequences derived from negative sense single strand RNA viruses is due, in part, to the presence of RNA splice sites in the subject RNAs. These viruses replicate cytoplasmically in the naturally occurring host cell environment. Accordingly, there is a lack of selective pressure against RNA sequences which comprise one or more such splice sites since the enzymes which splice eukaryotic cell RNA are generally only present in the nucleus. However, since the recombinant expression system of the present invention is based, in one embodiment, on the introduction into a eukaryotic host cell of a DNA molecule encoding the viral protein of interest, the requisite synthesis of DNA complementary to the naturally occurring viral RNA gene would consequently also result in copying of any splice sites present in the RNA. Transcription of these DNAs will occur in the nucleus of the eukaryotic host cell thereby exposing RNA transcribed from the subject DNA to the nuclear RNA splicing enzymes of the host cell.
- 15 In terms of optimising the naturally occurring viral protein encoding nucleotide sequence, the number of splice sites which are deleted in any given situation would depend on the object to be achieved. For example, if it is desired to produce the full length viral protein, then all splice sites occurring within the protein coding region of the encoding nucleic acid molecule should be deleted. However, if it is desired to produce only a fragment of the subject protein (for example, the F_{sol} portion of the F protein which, as hereinbefore defined, does not comprise the transmembrane and cytoplasmic regions of the F protein) then only the splice sites within that region need be removed.

25 Deletion of the subject splice sites is preferably achieved by substituting one or more nucleotides which comprise a splice site recognition sequence such that this sequence is no longer recognised by the relevant RNA splicing enzyme. It should be understood, however, that any other suitable method of mutating the splice site may be utilised within the context of the present invention.

- 30 The present invention is therefore preferably directed to a method of facilitating production of a protein or derivative thereof from a negative sense single stranded RNA virus, said

- 33 -

method comprising expressing in a mammalian host cell a nucleic acid molecule encoding said protein or derivative thereof, the nucleotide sequence of which nucleic acid molecule is optimised for expression by said mammalian host cell wherein said optimisation is codon optimisation and/or nucleotide splice site deletion.

5

Preferably, the subject protein is a fusion protein (more particularly the F protein), N, P or SH.

Preferably, the negative sense single stranded RNA virus is a virus of the family
10 Paramyxoviridae. More preferably the virus is of the sub-family Pneumovirinae and still more preferably the subject virus is a virus of the genus Pneumovirus. Most preferably, the virus is respiratory syncytial virus.

It should be understood that the present invention extends to the use of derivatives of the
15 optimised nucleic acid sequences.

Most preferably, said codon optimisation comprises modification of at least one A and/or T comprising codon to express G and C, respectively and said mammalian splice site deletion comprises deletion of at least one RNA splice site. To the extent that the nucleic
20 acid molecule which is introduced into the host cell is a DNA molecule, the subject deletion would relate to the region of the DNA molecule which would encode the RNA splice site.

By "derivatives" is meant nucleic acid sequences derived from single or multiple
25 nucleotide substitutions, deletions and/or additions including fusion with other nucleic acid molecules. In accordance with this definition, "derivative" therefore extends to sequences comprising any one or more of the optimised codons and/or optimised splice site regions of <400>5, <400>6, <400>556, <400>559 or <400>562.

30 Reference to a "derivative" of the subject nucleotide sequence should also be understood to extend to nucleotide sequences comprising nucleic acid substitutions, deletions or

- 34 -

additions other than for the purpose of optimising codons. For example, an optimised viral protein encoding nucleotide sequence may additionally comprises endonuclease restriction sites which are not expressed by the naturally occurring counterpart of the subject sequence. These may be incorporated to facilitate the generation of protein mutants. In
5 one preferred embodiment, for example, the subject nucleotide sequence derivative comprises one or more of the endonuclease restriction sites expressed in <400>3 or <400>4.

In terms of a most preferred embodiment of the present invention, <400>1 defines the
10 protein encoding region of the naturally occurring respiratory syncytial virus F protein. <400>3 defines the <400>1 sequence as modified to incorporate endonuclease restriction sites designed to facilitate the generation of protein recombinants. <400>5 defines the F protein encoding nucleotide sequence of <400>3 further modified to minimise the presence of regions which would encode RNA splice sites and to express optimised
15 codons. The amino acid sequence encoded by these nucleotide sequences is provided in <400>7.

Expression of <400>5 in accordance with the method of the present invention will be sought where production of the full length F protein is required. This may occur, for
20 example, where expression of a functional molecule is required for the performance of function based screening assays designed to detect F protein modulatory agents. However, in another embodiment, production of a portion only of the F protein may be desired. For example, production of the F_{sol} portion is particularly desirable for the purpose of 3 dimensional structural analysis, by X-ray crystallography, of the F protein active regions.
25 Furthermore, F_{sol} portion production facilitates the rational identification, modification and design of F protein modulatory agents based on analysing the agent in terms of its physical interaction with the F2 and F1 portions. In this regard, <400>2 defines the protein encoding region of the naturally occurring respiratory syncytial viral F_{sol} portion of the F protein. <400>4 defines the <400>2 sequence as modified to incorporate endonuclease
30 restriction sites designed to facilitate the generation of protein recombinants. <400>6 defines the F_{sol} protein encoding nucleotide sequence of <400>4 further modified to

- 35 -

minimise the presence of regions which would encode RNA splice sites and to express optimised codons. The amino acid sequence encoded by these nucleotide sequences is provided in <400>8.

- 5 According to this preferred embodiment there is provided a method of facilitating production of F protein or derivative thereof from respiratory syncytial virus, said method comprising expressing in a mammalian host cell a nucleic acid molecule encoding said F protein or derivative thereof, the nucleotide sequence of which nucleic acid molecule is optimised for expression by said mammalian host cell wherein said optimisation is
10 nucleotide splice site deletion.

- In another preferred embodiment the present invention is directed to a method of facilitating production of a F_{sol} portion of an F protein or derivative thereof from respiratory syncytial virus, said method comprising expressing in a host cell a nucleic acid
15 molecule encoding said F_{sol} portion or derivative thereof, the nucleotide sequence of which nucleic acid molecule is optimised for expression by said mammalian host cell wherein said optimisation is nucleotide splice site deletion.

- In still another preferred embodiment there is provided a method of facilitating production
20 of F protein or derivative thereof from respiratory syncytial virus, said method comprising expressing in a mammalian host cell a nucleic acid molecule encoding said F protein or derivative thereof, the nucleotide sequence of which nucleic acid molecule is optimised for expression by said mammalian host cell wherein said optimisation is codon optimisation.

- 25 In yet another preferred embodiment the present invention is directed to a method of facilitating production of a F_{sol} portion of an F protein or derivative thereof from respiratory syncytial virus, said method comprising expressing in a host cell a nucleic acid molecule encoding said F_{sol} portion or derivative thereof, the nucleotide sequence of which nucleic acid molecule is optimised for expression by said mammalian host cell wherein
30 said optimisation is codon optimisation.

- 36 -

In another preferred embodiment there is provided a method of facilitating production of F protein or derivative thereof from respiratory syncytial virus, said method comprising expressing in a mammalian host cell a nucleic acid molecule encoding said F protein or derivative thereof, the nucleotide sequence of which nucleic acid molecule is optimised for
5 expression by said mammalian host cell wherein said optimisation is nucleotide splice site deletion and codon optimisation.

In still yet a more preferred embodiment, there is provided a method of facilitating the production of a F protein or derivative thereof from a respiratory syncytial virus, said
10 method comprising expressing in a host cell the nucleotide sequence set forth in <400>5 or derivative thereof.

Preferably said nucleotide sequence is substantially as set forth in <400>5.

15 In another preferred embodiment, there is provided a method of facilitating the production of a F_{sol} portion of an F protein or derivative thereof from respiratory syncytial virus, said method comprising expressing in a host cell the nucleotide sequence set forth in <400>6 or derivative thereof.

20 Preferably said nucleotide sequence is substantially as set forth in <400>6.

In terms of another most preferred embodiment of the present invention, <400>555 defines the protein encoding region of the naturally occurring respiratory syncytial virus P protein. <400>556 defines the P protein encoding nucleotide sequence of <400>555 as modified to
25 express optimised codons. The amino acid sequence encoded by this nucleotide sequences is provided in <400>554.

According to this preferred embodiment there is provided a method of facilitating production of P protein or derivative thereof from respiratory syncytial virus, said method
30 comprising expressing in a mammalian host cell a nucleic acid molecule encoding said P protein or derivative thereof, the nucleotide sequence of which nucleic acid molecule is

- 37 -

optimised for expression by said mammalian host cell wherein said optimisation is codon optimisation.

5 In still a more preferred embodiment, there is provided a method of facilitating the production of a P protein or derivative thereof from respiratory syncytial virus, said method comprising expressing in a host cell the nucleotide sequence set forth in <400>556 or derivative thereof.

Preferably said nucleotide sequence is substantially as set forth in <400>556.

10

In terms of yet another most preferred embodiment of the present invention, <400>558 defines the protein encoding region of the naturally occurring respiratory syncytial virus N protein. <400>559 defines the N protein encoding nucleotide sequence of <400>558 as modified to express optimised codons. The amino acid sequence encoded by this
15 nucleotide sequence is provided in <400>557.

According to this preferred embodiment there is provided a method of facilitating production of N protein or derivative thereof from respiratory syncytial virus, said method comprising expressing in a mammalian host cell a nucleic acid molecule encoding said N
20 protein or derivative thereof, the nucleotide sequence of which nucleic acid molecule is optimised for expression by said mammalian host cell wherein said optimisation is codon optimisation.

In still a more preferred embodiment, there is provided a method of facilitating the
25 production of a N protein or derivative thereof from respiratory syncytial virus, said method comprising expressing in a host cell the nucleotide sequence set forth in <400>559 or derivative thereof.

Preferably said nucleotide sequence is substantially as set forth in <400>559.

30

- 38 -

In terms of still yet another most preferred embodiment of the present invention, <400>561 defines the protein encoding region of the naturally occurring respiratory syncytial virus SH protein. <400>562 defines the N protein encoding nucleotide sequence of <400>561 as modified to express optimised codons. The amino acid sequence encoded by this
5 nucleotide sequence is provided in <40>560.

According to this preferred embodiment there is provided a method of facilitating production of SH protein or derivative thereof from respiratory syncytial virus, said method comprising expressing in a mammalian host cell a nucleic acid molecule encoding
10 said SH protein or derivative thereof, the nucleotide sequence of which nucleic acid molecule is optimised for expression by said mammalian host cell wherein said optimisation is codon optimisation.

In still a more preferred embodiment, there is provided a method of facilitating the
15 production of a SH protein or derivative thereof from respiratory syncytial virus, said method comprising expressing in a host cell the nucleotide sequence set forth in <400>562 or derivative thereof.

Preferably said nucleotide sequence is substantially as set forth in <400>562.
20

In terms of performing the present invention, methods of deriving and recombinantly expressing nucleic acid molecules will be well known to those of skill in the art as will methodology directed to adding, deleting and/or substituting nucleic acids in a given nucleotide sequence.
25

In another aspect, the present invention should be understood to extend to the optimised nucleic acid molecules described herein and to the expression products derived therefrom.

In yet another aspect, the inventors have surprisingly determined that induction of F
30 protein functional activity requires not one but two proteolytic cleavage events. The occurrence of these two cleavage events results in the excision of a peptide region from the

- 39 -

non-fully functional F protein. Prior to the advent of the present invention, it was thought that F protein activation was the result of a single cleavage event which occurred at the F2/F1 junction. Without limiting the invention to any one theory or mode of action, it is thought that the F2 portion of the non-fully functional F protein in fact comprises an
5 intervening sequence of amino acids which spans the region between the newly identified cleavage site and the F2/F1 junction and which is excised in order to facilitate formation of a functional F glycoprotein. This intervening peptide sequence is thought to comprise excess amino acids and up to three glycosylation sites depending on the particular virus strain from which the F protein is derived. Down-regulation or other form of interference
10 with cleavage at the newly identified cleavage site would therefore interfere with the induction of F protein functional activity.

Accordingly, another aspect of the present invention is directed to a method of regulating the functional activity of a viral F protein, which protein in its non-fully functional form
15 comprises an F2 portion linked, bound or otherwise associated with an F1 portion, which F2 portion comprises an intervening peptide sequence, said method comprising modulating cleavage of said intervening peptide sequence wherein excision of at least part of said intervening sequence from said non-fully functional form of said F protein up-regulates F protein functional activity.

20

Reference to the subject F protein being in a "non-fully functional form" should be understood to mean that the subject F protein exhibits either no functional activity or a lesser degree of functional activity than the fully cleaved F protein, that is, the F protein which has undergone *both* cleavage events. Accordingly, "up-regulation" of F protein
25 functional activity should be understood to refer to the induction of a degree or range of functional activities greater than that exhibited by the subject F protein in its non-fully cleaved form. In its natural environment, all F proteins are synthesised in a form which comprises a F2 portion located proximally to a F1 portion. The F1 portion region of the F protein comprises a transmembrane region and an intracellular domain (Collins *et al*,
30 1996). Reference to a "non-fully functional form" of the F protein should also be understood to extend to forms of the F protein which have undergone only partial cleavage.

- 40 -

For example, the subject non-fully functional form of the F protein may only have undergone cleavage of the previously known cleavage site but not yet at the newly identified cleavage site.

5 Prior to the advent of the present invention, it was thought that activation of the F protein occurred following cleavage at the F protein site defined by the sequence KKRKRR (<400>563) thereby cleaving the F2 portion of the non-fully functional F protein from the F1 portion. The F1 portion of the F protein is defined, in Figure 1, as commencing at the F residue which follows the cleavage recognition site KKRKRR. However, the precise
10 location at which this cleavage event occurs is not actually known. Accordingly, it should be understood that the cleavage event may occur either between two residues located proximally to the cleavage recognition site KKRKRR or it may occur between two residues within this site. The definitions of "F2 portion", "F1 portion" and "F2/F1 junction" as provided herein should therefore be interpreted in light of this understanding.

15

As detailed above, the inventors have now determined that cleavage at this region alone will not fully activate the F protein. Rather, a second cleavage event must occur at an F protein site distinct from that of the known cleavage site (the known cleavage site being referred to as "site 1"). This second cleavage site is located in the amino terminus
20 direction of the previously known cleavage site and is characterised by expression of the cleavage recognition sequence RARR (<400>564) (herein referred to as "site 2"). When considered in light of the structure of the F protein as it was previously understood (and as depicted in Figure 1) site 2 is located within the F2 portion of the F protein while the previously known cleavage site is located at the F2/F1 portion junction.

25

For the purpose of the present invention, it should be understood that the F protein amino acid sequence located in the amino terminus direction of cleavage site 1 is herein referred to as the F2 portion while the amino acid sequence located in the carboxy terminus direction of the cleavage site 1 is herein referred to as the F1 portion. The newly identified
30 cleavage site is therefore located within the F2 portion. The F protein amino acid sequence located between the site 1 and site 2 points of cleavage is herein referred to as the

- 41 -

"intervening sequence". Accordingly, in light of the definition herein provided, the "intervening sequence" forms part of the F2 portion of the non-fully functional form of the F protein. Excision of "at least part of" said intervening sequence should be understood to mean that at least a portion of the sequence which is excised following the two cleavage
5 events is derived from the intervening sequence region as herein defined. However, it should be understood that the excised sequence may also comprise part of the non-intervening sequence region of the F2 and/or F1 portion sequences as herein defined.

Without limiting the present invention to any one theory or mode of action, it is thought
10 that cleavage of the intervening sequence at the two cleavage sites results in complete disassociation of the intervening sequence from the F protein. Accordingly, the term "excision" is intended to encompass complete disassociation of the intervening sequence from the non-fully functional form of the F protein in order to form the functionally active F protein. However this term should also be understood to extend to a cleavage event
15 which does not necessarily result in complete disassociation of at least part of the intervening sequence but leads to a conformational change in the secondary or tertiary structure of the intervening sequence and/or the F2/F1 portions. For example, in some circumstances an appropriate conformational shift in the intervening sequence relative to the F2 and F1 portions may be sufficient to achieve some up-regulation of the functional
20 activity of the F protein. It should also be understood that the two cleavage events may occur concurrently in order to effect excision. Alternatively, the cleavage events may occur consecutively. For example, cleavage at site 1 may occur initially, followed by cleavage at site 2 (and hence formation of the fully functional form of the F protein) at a subsequent point in time. The present invention should also be understood to extend to a
25 sequence of cleavage events commencing with cleavage at site 2.

The present invention is exemplified with respect to respiratory syncytial virus F protein. The respiratory syncytial virus F protein amino acid sequence is defined <400>7. In accordance with the amino acid sequence numbering provided in <400>7, the previously
30 known cleavage site is located at the region of the F protein defined by the amino acid sequence KKRKRR, which sequence spans amino acid numbers 131 to 136 of <400>7.

- 42 -

The second cleavage point, which has been identified by the present inventors, is localised to the region of the F protein defined by the amino acid sequence RARR, which sequence spans amino acid numbers 106-109 of <400>7.

- 5 In a preferred embodiment the present invention is directed to a method of regulating the functional activity of a Paramyxoviridae derived F protein, which protein in its non-fully functional form comprises an F2 portion linked, bound or otherwise associated with an F1 portion, which F2 portion comprises an intervening peptide sequence, said method comprising modulating cleavage of said intervening peptide sequence wherein excision of
10 at least part of said intervening sequence from said non-fully functional form of said F protein up-regulates F protein functional activity.

Still more preferably said F protein is derived from the Genus Pneumovirus and still more preferably said virus is respiratory syncytial virus.

15

- In a most preferred embodiment there is provided a method of regulating the functional activity of a respiratory syncytial virus F protein, which protein in its non-fully functional form comprises an F2 portion linked, bound or otherwise associated with an F1 portion, which F2 portion comprises an intervening peptide sequence, said method comprising
20 modulating cleavage of said intervening peptide sequence, wherein excision of at least part of said intervening sequence from said non-fully functional form of said F protein up-regulates F protein functional activity and wherein said cleavage events occur at the cleavage sites defined by the peptide sequences RARR (<400>564) and KKRKRR (<400>563).

25

- That the subject cleavage events "occur at" a given cleavage site should be understood to mean that cleavage of the F protein amino acid sequence will involve cleavage of the bonding mechanism associated with anyone or more of the amino acids comprising the defined sites. Without limiting the invention in any way, the amino acids comprising the
30 cleavage sites define the peptide sequence recognised by the proteolytic enzyme which cleaves the subject F protein (Steiner, 1998).

- 43 -

In a related aspect, the present invention provides a method of regulating the functional activity of a viral F protein, which protein in its non-fully functional form comprises the structure:

5 X₁, X₂, X₃

wherein:

 X₁ comprises the non-intervening peptide sequence region of the F₂ portion;

 X₂ comprises the intervening peptide sequence region of the F₂ portion; and

10 X₃ comprises the F₁ portion

said method comprising modulating cleavage of said intervening peptide sequence wherein excision of at least part of said intervening sequence from said non-fully functional form of said F protein up-regulates F protein functional activity.

15

The representation X₁, X₂, X₃ is not to be taken as imposing any sequential constraints on the subject F protein and the present invention encompasses any conformational secondary and/or tertiary structural arrangement of X₁, X₂, X₃ to the extent that X₁ and X₃ are both linked, bound or otherwise associated with X₂ in the subject F protein's non-fully
20 functional form.

Reference to the "non-intervening peptide sequence region" of F₂ should be understood as a reference to that part of the F₂ subunit which does not form part of the intervening sequence as herein defined.

25

Preferably said virus is a virus from the family Paramyxoviridae and still more preferably is a virus of the Genus Pneumovirus. Most preferably said virus is respiratory syncytial virus.

- 44 -

In another preferred embodiment said cleavage events occur at the cleavage sites comprising X_2 and defined by the peptide sequences RARR (<400>564) and KKRKRR (<400>563).

5 Modulating cleavage of the intervening sequence can be achieved by any one of a number of methods including, but in no way limited to:

(i) Contacting the F protein or F protein encoding nucleic acid molecule with a
proteinaceous or non-proteinaceous molecule (herein referred to as an "agent")
10 which up-regulates or down-regulates cleavage of either one or both of the
cleavage sites comprising the intervening sequence. The proteinaceous or non-
proteinaceous molecule may achieve this objective by functioning as either an
agonist or antagonist of the cleavage event, for example. This molecule may act in
any one of a number of ways including interacting with the subject F protein or
15 interacting with the enzymes which recognise the cleavage sites comprising the F
protein.

(ii) Mutating the amino acid sequence of the F protein cleavage site such that
proteolytic cleavage cannot occur. This can be performed at either the amino acid
20 sequence level (for example by adding, substituting or deleting an amino acid in the
newly identified cleavage site) or at the nucleotide level such that the transcribed
and translated F protein expression product does not express a functional form of
the subject cleavage site.

25 Said proteinaceous molecule may be derived from natural or recombinant sources
including fusion proteins or following, for example, natural product screening. Said non-
proteinaceous molecule may be, for example, a nucleic acid molecule or may be derived
from natural sources, such as for example natural product screening or may be chemically
synthesised. The present invention contemplates chemical analogues of the F protein
30 capable of acting as agonists or antagonists of either the fully functional or non-fully
functional F protein. Chemical agonists may not necessarily be derived from the F protein

- 45 -

but may share certain conformational similarities. Alternatively, chemical agonists may be specifically designed to mimic certain physiochemical properties of the F protein. Antagonists may be any compound capable of blocking, inhibiting or otherwise preventing F protein from carrying out its normal biological function. Antagonists include
5 monoclonal antibodies specific for the F protein, or parts of the F protein, and antisense nucleic acids which prevent transcription and/or translation of the F protein encoding nucleic acid molecule or mRNA in mammalian cells.

Although the preferred method is to inhibit, retard or otherwise down-regulate F protein
10 functional activity by preventing cleavage of the non-fully functional F protein form and subsequent activation, up-regulation of F protein functional activity may be desired in certain circumstances. In this regard, use of agonistic agents which augment and/or induce the cleavage events herein described may be utilised. Reference to "down-regulating" F protein functional activity should be understood to encompass prevention of the functional
15 activation of the non-fully functional F protein.

Accordingly, in a most preferred embodiment there is provided a method of inhibiting, retarding or otherwise down-regulating the functional activity of a Paramyxoviridae derived F protein, which protein in its non-fully functional form comprises an F2 portion
20 linked, bound or otherwise associated with an F1 portion, which F2 portion comprises an intervening peptide sequence, said method comprising inhibiting or otherwise down-regulating cleavage of said intervening peptide sequence.

Preferably said F protein is derived from the Genus Pneumovirus and still more preferably
25 said virus is respiratory syncytial virus. Most preferably said cleavage events occur at the cleavage sites defined by peptide sequences RARR (<400>564) and KKRKRR (<400>563).

In another most preferred embodiment the present invention provides a method of down-
30 regulating the functional activity of a Paramyxoviridae derived F protein, which protein in its non-fully functional form comprises the structure:

- 46 -

 $X_1X_2X_3$

wherein:

- 5 X_1 comprises the non-intervening peptide sequence region of the F2 portion;
 X_2 comprises the intervening peptide sequence region of the F2 portion; and
 X_3 comprises the F1 portion

 said method comprising inhibiting or otherwise down-regulating cleavage of said
10 intervening peptide sequence.

 Preferably said F protein is derived from the Genus Pneumovirus and still more preferably
 said virus is respiratory syncytial virus. Most preferably said cleavage events occur at the
 cleavage sites defined by peptide sequences RARR (<400>564) and KKRKRR
15 (<400>563).

 Without limiting the present invention to any one theory or mode of action, the F proteins
 of viruses of the family Paramyxoviridae are involved in facilitating fusion between the
 virus envelope and the host cell plasma membrane in order to effect infection. Further, it
20 is thought that the F proteins are also inserted into the host plasma membrane where,
 during maturation, the virions bud through the region of the membrane containing these
 proteins. Accordingly, it is thought that down-regulating F protein functional activity will
 inhibit or otherwise reduce virion fusion with and infection of a potential host cell and/or
 virion budding. Accordingly, the development of a method for recombinantly expressing
25 the F protein by eukaryotic cells, and in particular mammalian cells, now facilitates the
 development of screening assays, utilising the F protein produced in accordance with the
 method of the present invention, for the purpose of identifying agents capable of
 modulating F protein functional activity, and preferably, down-regulating F protein
 functional activity.

30

- 47 -

Screening for agents which modulate F protein functional activity can be achieved by any one of a number of suitable methods, which would be known to those of skill in the art, including but not limited to:

- 5 (i) High throughput screening for agents which modulate F protein functional activities utilising assays based on the detection of changes in F protein functioning. Such changes may be detected directly or indirectly.

10 An example of indirect detection of modulation of F protein functioning includes the screening of agents on cultured cells which have been co-transfected with the F protein encoding nucleic acid molecule of the present invention and a virus which utilises the F protein in order to propagate. In this regard, either the full length F protein encoding nucleic acid sequence can be utilised or a partial sequence which encodes a functionally active F protein portion can be used. By assessing cell
15 viability it can be determined whether the subject agent inhibits or down-regulates F protein functioning thereby preventing F protein mediated propagation of cell to cell fusion. This would be evident by continued cell viability. A typical assay of this type can be performed, for example, in 293 cells which have been transiently co-transfected with a plasmid encoding the adenoviral VA RNA genes.

20

(ii) Antibody Recognition Assays

The use of antibodies which bind to conformational epitopes is a recognised method for assessing whether a protein's three dimensional structure differs from
25 the natural state. Thus an assay can be conducted on protein exposed to agents that are expected to modulate function via perturbation of the native conformation or interference with a functional conformational transition. A number of suitable F-specific antibodies and their target sites have been identified by workers in the field (see for example Lopez et al., 1998 and references therein). For example, F protein
30 exposed to agents intended to modulate F function is subsequently incubated with F specific monoclonal antibodies using an ELISA format. Reduction or increase in F binding relative to F which has not been exposed to agents is measured by addition

- 48 -

of polyclonal antibody to F followed by suitable detection reagents according to standard methods.

(iii) Immunisation leading to protection and/or virus neutralisation

5 RSV is known to infect a wide range of animal species when inoculated experimentally into the respiratory tract and several small animal experimental models have been described (see for example Collins et al., 1996 and references therein). These models can be used to determine whether immunisation is
10 protective and/or results in the production of a virus neutralising response.

An example of a suitable method is as follows: Cotton rats (average weight 100 g) are anesthetized with methoxyflurane and a sample of pre-immune blood harvested via standard procedures. While anesthetized, the cotton rats are administered a
15 suitable quantity of agent (for example, purified F protein) via an appropriate route (for example, intramuscular injection or intranasal instillation). The cotton rats are housed for an appropriate period (generally several days to weeks depending on the agents under consideration and the objectives of the study) and then anesthetized as above. Anesthetized animals are bled to obtain a post-immunization sample and
20 infected with 100,000 plaque forming units of a suitable RSV strain (for example, RSV Long). Four days later the animals are sacrificed and lungs harvested aseptically. Protective efficacy of the agent is measured by determination of the effect on whole lung virus titre. Briefly, lungs are homogenised in sterile saline (1:10 w/v) and virus concentration determined by standard methods (for example,
25 plaque assay).

To determine whether the agent elicited a neutralising response, pre-immunization and post-immunization samples and control samples are examined using a virus neutralization test. An example of such a test is as follows. Sera are prepared from
30 the blood samples according to standard methods. Serial dilutions of the sera are then prepared and mixed with a known concentration of RSV (for example, 100 plaque forming units of RSV Long). Mixtures are incubated for 1 hour at room temperature before being assayed for virus concentration by standard methods (for

- 49 -

example, plaque assay). A neutralizing response is characterised by reduction in virus titre in comparison to control samples.

Accordingly, in another aspect there is provided a method for detecting an agent capable of
5 regulating the functional activity of a viral F protein or derivative thereof said method comprising contacting a eukaryotic cell expressing an optimised nucleic acid molecule encoding said viral F protein or derivative thereof, as hereinbefore described, with a putative modulatory agent and detecting an altered expression phenotype and/or functional activity.

10

It should be understood that the subject agent may act via any mechanism including, but not limited to, modulating the cleavage events hereinbefore described.

In yet another aspect there is provided a method for detecting an agent capable of
15 regulating the functional activity of a viral F protein or derivative thereof said method comprising contacting a host cell, which host cell expresses a nucleic acid molecule encoding the non-fully functional form of said viral F protein or derivative thereof as hereinbefore described, with a putative modulatory agent and detecting an altered expression phenotype and/or altered functional activity wherein said agent modulates
20 cleavage of the intervening peptide sequence.

To the extent that this aspect of the present invention is directed to screening for agents which modulate the site 2 cleavage event, it should be understood that this methodology is not limited to systems expressing an optimised nucleic acid sequence but extends to
25 systems utilising any method of expressing the subject F protein.

Reference to a "modulatory agent" should be understood as a reference to an agent which down-regulates, up regulates or otherwise alters at least one functional activity of the subject F protein. For example, the agent may increase or decrease the level of activity of
30 the F protein or it may entirely inhibit F protein functioning. Although the preferred method is to identify agents which inhibit F protein functional activity, for example by

- 50 -

preventing cleavage of the non-fully functional form of the F protein, thereby providing a potential antiviral therapy, the identification of agents which up regulate F protein functional activity may be desired in certain circumstances. For example, it is thought that an agent which causes the F protein to prematurely initiate the conformational changes required for fusion would be inactivating.

Still more preferably, said viral F protein is a Pneumovirus F protein and yet still more preferably a respiratory syncytial virus F protein. Most preferably, said codon optimised nucleic acid molecule is the nucleic acid molecule defined in <400>5.

Preferably, said regulation is inhibition, retardation or other form of down-regulation.

Reference to "functional activity" should be understood as a reference to any one or more of the functions which the F protein performs. Accordingly, an agent which modulates the functional activity of the F protein may modulate all or only some of the functions which the F protein performs. The phrase "functional activity" should be understood to include within its scope the cleavage events which the F protein undergoes.

In addition to screening for agents which modulate F protein functional activity utilising function based assays of the type described above, the development of methodology which facilitates production of the F protein or derivatives thereof also facilitates the screening, analysis, rational design and/or modification of agents for modulating F protein functional activity based on analysis of the physical interaction of a putative agent or lead compound with the subject F protein or derivative thereof.

Specifically, *in vitro* production of the F protein or derivative thereof, which is now possible in light of the development of the present invention, now facilitates analysis of the tertiary structure of the F protein by techniques such as X-ray crystallography. Of particular value in this regard is the fact that the present invention permits production of useful quantities of the F protein F_{sol} portion.

- 51 -

Accordingly, another aspect of the present invention is directed to a method for analysing, designing and/or modifying an agent capable of interacting with a viral F protein or derivative thereof and modulating at least one functional activity associated with said protein, which protein is produced in accordance with the method of the present invention
5 said method comprising contacting said F protein or derivative thereof with a putative agent and assessing the degree of interactive complementarity of said agent with said protein.

Preferably said viral F protein is a Pneumovirus F protein and even more preferably the F_{sol} portion of said Pneumovirus F protein. Still more preferably, said F_{sol} portion is defined by
10 the amino acid sequence of <400>8.

It should be understood that the F protein which is contacted with the putative agent for evaluation of interactive complementarity may be recombinantly produced. However, it should also be understood that the subject protein may take the form of an image based on
15 the structure elucidated via analysis of the F protein produced in accordance with the method of the present invention, such as an electron density map, molecular models (including, but not limited to, stick, ball and stick, space filling or surface representation models) or other digital or non-digital surface representation models or image, which facilitates the analysis of F protein: agent interactions utilising techniques and software
20 which would be known to those of skill in the art. For example, interaction analyses can be performed utilising techniques such as Biacore real-time analysis of on and off-rates and dissociation constants for binding of ligands (Gardsvoll *et al*, 1999; Hoyer-Hansen *et al*, 1997; Ploug, 1998; Ploug *et al*, 1994; 1995; 1998) and NMR perturbation studies (Stephens *et al*, 1992).

25 Reference to "assessing the degree of interactive complementarity" of an agent with the subject F protein should be understood as a reference to elucidating any feature of interest including, but not limited to, the nature and/or degree of interaction between the subject F protein and an agent of interest. As detailed above, any suitable technique can be utilised.
30 Such techniques would be known to the person of skill in the art and can be utilised in this regard. In terms of the nature of the subject interaction, it may be desirable to assess the

- 52 -

- types of interactive mechanisms which occur between specific residues of any given agent and those of the F protein (for example, peptide bonding or formation of hydrogen bonds, ionic bonds, van der Waals forces, etc.) and/or their relative strengths. It may also be desirable to assess the degree of interaction which occurs between an agent of interest and the subject F protein. For example, by analysing the location of actual sites of interaction between the subject agent and F protein it is possible to determine the quality of fit of the agent into any region of the F protein and the relative strength and stability of that binding interaction. For example, if it is the object that F protein functioning be blocked, an agent which interacts with the F protein such that it blocks or otherwise hinders (for example, sterically hinders or chemically or electrostatically repels) F2/F1 cleavage will be sought. The form of association which is required in relation to modulating F protein functioning may not involve the formation of any interactive bonding mechanism, as this is traditionally understood, but may involve a non-bonding mechanism such as the proximal location of a region of the agent relative to the subject binding region of the F protein, for example, to effect steric hindrance with respect to the binding of an activating molecule. Where the interaction takes the form of hindrance or the creation of other repulsive forces, this should nevertheless be understood as a form of "interaction" despite the lack of formation of any of the traditional forms of bonding mechanisms.
- It should also be understood that the F protein which is utilised either in a physical form or as an image, as hereinbefore discussed, to assess the interactive complementarity of a putative agent may be a naturally occurring form of the F protein or it may be a derivative, homologue, analogue, mutant, fragment or equivalent thereof. The derivative, homologue, analogue, mutant, fragment or equivalent thereof may take either a physical or non-physical (such as an image) form.

The determination of F protein binding regions has been made possible only by development of the present invention which has permitted F protein production and thereby has facilitated determination of the three dimensional structure of the F protein and the identification and/or rational modification and design of agents which can be used to modulate F protein functioning.

- 53 -

Without limiting the application of the present invention in any way, the method of the present invention facilitates the analysis, design and/or modification of agents capable of interacting with the F protein. In this regard, reference to "analysis, design and/or modification" of an agent should be understood in its broadest sense to include:

- 5
- (i) Randomly screening (for example, utilising routine high-throughput screening technology) to identify agents which exhibit some modulatory capacity with respect to F protein functional activity and then analysing the precise nature and magnitude of the agent's modulatory capacity utilising the method of this aspect of
- 10 the present invention. In this regard, existing crystals could be soaked with said agents or co-crystallisation could be performed. A combination of modelling and synthetic modification of the local compound together with mutagenesis of the F protein could then be performed for example. In screening for agents which may modulate activity, standard methods of phage display and also combinatorial chemistry may be utilised (Goodson *et al.*, 1994; Terrett., 2000). Such interaction
- 15 studies can also be furthered utilising techniques such as the Biacore analysis and NMR perturbation studies. Such agents are often commonly referred to as "lead" agents in terms of the random screening of proteinaceous or non-proteinaceous molecules for their capacity to function either agonistically or antagonistically.
- 20 Further, for example, binding affinity and specificity could be enhanced by modifying lead agents to maximise interactions with the F protein. Such analyses would facilitate the selection of agents which are the most suitable for a given purpose. In this way, the selection step is based not only on *in vitro* data but also on a technical analysis of sites of agent: F protein interaction in terms of their
- 25 frequency, stability and suitability for a given purpose. For example, such analysis may reveal that what appears to be an acceptable *in vitro* activity in respect of a randomly identified agent is in fact induced by a highly unstable interaction due to the presence of proximally located agent: F protein sites which exhibit significant repulsive forces thereby de-stabilising the overall interaction between the agent and
- 30 the F protein. This would then facilitate the selection of another prospective lead compound, exhibiting an equivalent degree of *in vitro* activity, but which agent

- 54 -

does not, upon further analysis, involve the existence of such de-stabilising repulsive forces.

5 Screening for the modulatory agents herein defined can be achieved by any one of several suitable methods, including in silico methods, which would be well known to those of skill in the art and which are, for example, routinely used to randomly screen proteinaceous and non-proteinaceous molecules for the purpose of identifying lead compounds.

10 These methods provide a mechanism for performing high throughput screening of putative modulatory agents such as the proteinaceous or non-proteinaceous agents comprising synthetic, recombinant, chemical and natural libraries.

(ii) 15 The candidate or lead agent (for example, the agent identified in accordance with the methodology described in relation to point (i)) could be modified in order to maximise desired interactions (for example, binding affinity to specificity) with the F protein and to minimise undesirable interactions (such as repulsive or otherwise de-stabilising interactions). Such modification is only possible in light of knowledge of the three-dimensional structure of the F protein and the capacity 20 therefore to identify regions of functional importance, thereby facilitating the structural modification of an agent to maximise an agonistic or antagonistic interaction. Such methodology is particularly applicable to rational drug design.

25 Methods of modification of a candidate or lead agent in accordance with the purpose as defined herein would be well known to those of skill in the art. For example, a molecular replacement program such as Amore (Navaza, 1994) may be utilised in this regard. The method of the present invention also facilitates the mutagenesis of known signal inducing agents in order to ablate or improve signalling activity.

30

- 55 -

- (iii) In addition to analysing fit and/or structurally modifying existing molecules, the method of the present invention also facilitates the rational design and synthesis of an agent, such as an agonistic or antagonistic agent, based on theoretically modelling an agent exhibiting the desired F protein interactive structural features followed by the synthesis and testing of the subject agent.

It should be understood that any one or more of applications (i) – (iii) above, may be utilised in identifying a particular agent.

- 10 In a related aspect, the present invention should be understood to extend to the agents identified utilising any of the methods hereinbefore defined. In this regard, reference to an agent should be understood as a reference to any proteinaceous or non-proteinaceous molecule which modulates at least one F protein functional activity. As hereinbefore discussed, to the extent that the present invention encompasses methods of screening for agents utilising F proteins produced in accordance with the expression system hereinbefore defined, this is not to be taken as a restriction on the methodology which is employed to screen for agents which modulate the newly identified cleavage event. In this regard, the present invention extends to agents identified utilising F protein molecules or derivatives thereof howsoever produced.

20

Accordingly, the present invention is directed to an agent capable of interacting with a viral F protein and modulating at least one functional activity associated with said viral protein.

- 25 Preferably, said agent is identified in accordance with the methods hereinbefore defined.

More preferably, said agent is an antagonist which interacts with a sequence selected from:

CFASGQNITE <400>9
 ASGQNITEEF <400>11
 GQNITEEFYQ <400>13
 NITEEFYQST <400>15
 TEEFYQSTCS <400>17

FASGQNITEE <400>10
 SGQNITEEFY <400>12
 QNITEEFYQS <400>14
 ITEEFYQSTC <400>16
 EEFYQSTCSA <400>18

- 56 -

EFYQSTCSAV <400>19	FYQSTCSAVS <400>20
YQSTCSAVSK <400>21	QSTCSAVSKG <400>22
STCSAVSKGY <400>23	TCSAVSKGYL <400>24
CSAVSKGYLS <400>25	SAVSKGYLSA <400>26
AVSKGYLSAL <400>27	VSKGYLSALR <400>28
SKGYLSALRT <400>29	KGYSALALRTG <400>30
GYLSALRTGW <400>31	YLSALRTGWY <400>32
LSALRTGWYT <400>33	SALRTGWYTS <400>34
ALRTGWYTSV <400>35	LRTGWYTSVI <400>36
RTGWYTSVIT <400>37	TGWYTSVITI <400>38
GWYTSVITIE <400>39	WYTSVITIEL <400>40
YTSVITIELS <400>41	TSVITIELSN <400>42
SVITIELSNI <400>43	VITIELSNIK <400>44
ITIELSNIKK <400>45	TIELSNIKKN <400>46
IELSNIKKNK <400>47	ELSNIKKKNC <400>48
LSNIKKKNCN <400>49	SNIKKKNKCNG <400>50
NIKKKNKCNGT <400>51	IKKNKCNGTD <400>52
KKNKCNGTDA <400>53	KNKCNGTDAK <400>54
NKCNGTDAKV <400>55	KCNGTDAKVK <400>56
CNGTDAKVKL <400>57	NGTDAKVKLI <400>58
GTDAAKVLIK <400>59	TDAAKVLIKQ <400>60
DAKVLIKQOE <400>61	AKVLIKQOEL <400>62
KVKLIKQELD <400>63	VKLIKQELDK <400>64
KLIKQELDKY <400>65	LIKQELDKYK <400>66
IKQELDKYKN <400>67	KQELDKYKNA <400>68
QELDKYKNAV <400>69	ELDKYKNAVT <400>70
LDKYKNAVTE <400>71	DKYKNAVTEL <400>72
KYKNAVTELQ <400>73	YKNAVTELQL <400>74
KNAVTELQLL <400>75	NAVTELQLLM <400>76
AVTELQLLMQ <400>77	VTELQLLMQS <400>78
TELQLLMQST <400>79	ELQLLMQSTQ <400>80
LQLLMQSTQA <400>81	QLLMQSTQAT <400>82
LLMQSTQATN <400>83	LMQSTQATNN <400>84
MQSTQATNNR <400>85	QSTQATNNRA <400>86
STQATNNRAR <400>87	TQATNNRARR <400>88
QATNNRARRE <400>89	ATNNRARREL <400>90
TNNRARRELP <400>91	NNRARRELPR <400>92
NRARRELPRF <400>93	RARRELPRFM <400>94
ARRELPRFMN <400>95	RRELPRFMNY <400>96
RELPRFMNYT <400>97	ELPRFMNYTL <400>98
LPRFMNYTLN <400>99	PRFMNYTLNN <400>100
RFMNYTLNNA <400>101	FMNYTLNNAK <400>102
MNYTLNNAKK <400>103	NYTLNNAKKT <400>104
YTLNNAKKTN <400>105	TLNNAKKTNV <400>106
LNNAKKTNVT <400>107	NNAKKTNVTL <400>108
NAKKTNVTLS <400>109	AKKTNVTL SK <400>110
KKTNVTL SKK <400>111	KTNVTLSKRR <400>112
TNVTL SKKRK <400>113	NVTLSKRRKR <400>114
VTLSKRRKRR <400>115	TL SKRRKR RF <400>116
LSKRRKR RFL <400>117	SKRRKR RFLG <400>118
KKRRKR RFLGF <400>119	KRRKR RFLGFL <400>120
RKR RFLGFL L <400>121	KRRFLGFL LG <400>122
RRFLGFL LGV <400>123	RFLGFL LGVG <400>124
FLGFL LGVGS <400>125	LGFL LGVGS A <400>126

- 57 -

GFLLGVGSAI <400>127	FLLGVGSAIA <400>128
LLGVGSAIAS <400>129	LGVGSAIASG <400>130
GVGSAIASGV <400>131	VGSAIASGVA <400>132
GSAIASGVAV <400>133	SAIASGVAVS <400>134
AIASGVAVSK <400>135	IASGVAVSKV <400>136
ASGVAVSKVL <400>137	SGVAVSKVLH <400>138
GVAVSKVLHL <400>139	VAVSKVLHLE <400>140
AVSKVLHLEG <400>141	VSKVLHLEGE <400>142
SKVLHLEGEV <400>143	KVLHLEGEVN <400>144
VLHLEGEVNK <400>145	LHLEGEVNKI <400>146
HLEGEVNKIK <400>147	LEGEVNKIKS <400>148
EGEVNLIKSA <400>149	GEVNLIKSAI <400>150
EVNLIKSAI <400>151	VNLIKSAIIS <400>152
NLIKSAIIST <400>153	KIKSAIISTN <400>154
IKSAIISTNK <400>155	KSAIISTNKA <400>156
SAIISTNKAV <400>157	ALLSTNKAVV <400>158
LLSTNKAVVS <400>159	LSTNKAVVSL <400>160
STNKAVVSLS <400>161	TNKAVVSLSN <400>162
NKAVVSLSNG <400>163	KAVVSLSNGV <400>164
AVVSLSNGVS <400>165	VVSLSNGVSV <400>166
VSLSNGVSVL <400>167	SLSNGVSVLT <400>168
LSNGVSVLTS <400>169	SNGVSVLTSK <400>170
NGVSVLTSKV <400>171	GVSVLTSKVL <400>172
VSVLTSKVLD <400>173	SVLTSKVLDL <400>174
VLTSKVLDLK <400>175	LTSKVLDLKN <400>176
TSKVLDLKNY <400>177	SKVLDLKNYI <400>178
KVLDLKNYID <400>179	VLDLKNYIDK <400>180
LDLKNYIDKQ <400>181	DLKNYIDKQL <400>182
LKNYIDKQLL <400>183	KNYIDKQLLP <400>184
NYIDKQLLPI <400>185	YIDKQLLPIV <400>186
IDKQLLPIVN <400>187	DKQLLPIVKN <400>188
KQLLPIVKNQ <400>189	QLLPIVKNQS <400>190
LLPIVKNQSC <400>191	LPIVKNQSCS <400>192
PIVKNQSCSI <400>193	IVKNQSCSIS <400>194
VKNQSCSISN <400>195	NKQSCSISNI <400>196
KQSCSISNIE <400>197	QSCSISNIET <400>198
SCSISNIETV <400>199	CSISNIETVI <400>200
SISNIETVIE <400>201	ISNIETVIEF <400>202
SNIETVIEFQ <400>203	NIETVIEFQQ <400>204
IETVIEFQQK <400>205	ETVIEFQQKN <400>206
TVIEFQQKNN <400>207	VIEFQQKNNR <400>208
IEFQQKNNRL <400>209	EFQQKNNRLL <400>210
FQQKNNRLLE <400>211	QQKNNRLLEI <400>212
QKNNRLLEIT <400>213	KNNRLLEITR <400>214
NNRLLEITRE <400>215	NRLLEITREF <400>216
RLLEITREFS <400>217	LLEITREFSV <400>218
LEITREFSVN <400>219	EITREFSVNA <400>220
ITREFSVNAG <400>221	TREFSVNAGV <400>222
REFSVNAGVT <400>223	EFSVNAGVTT <400>224
FSVNAGVTPV <400>225	SVNAGVTPV <400>226
VNAGVTPVST <400>227	NAGVTPVST <400>228
AGVTPVSTYM <400>229	GVTPVSTYMT <400>230
VTPVSTYMLT <400>231	TPPVSTYMLT <400>232
TPVSTYMLTN <400>233	PVSTYMLTNS <400>234

- 58 -

VSTYMLTNSE	<400>235	STYMLTNSEL	<400>236
TYMLTNSELL	<400>237	YMLTNSELLS	<400>238
MLTNSELLSL	<400>239	LTNSELLSLI	<400>240
TNSELLSLIN	<400>241	NSSELLSLIND	<400>242
SELLSLINDM	<400>243	ELLSLINDMP	<400>244
LLSLINDMPI	<400>245	LSLINDMPIT	<400>246
SLINDMPITN	<400>247	LINDMPITND	<400>248
INDMPITNDQ	<400>249	NDMPITNDQK	<400>250
DMPITNDQKK	<400>251	MPITNDQKKL	<400>252
PITNDQKKLM	<400>253	ITNDQKKLMS	<400>254
TNDQKKLMSN	<400>255	NDQKKLMSNN	<400>256
DQKKLMSNNV	<400>257	QKKLMSNNVQ	<400>258
KKLMSNNVQI	<400>259	KLMSNNVQIV	<400>260
LMSNNVQIVR	<400>261	MSNNVQIVRQ	<400>262
SNNVQIVRQQ	<400>263	NNVQIVRQQS	<400>264
NVQIVRQQSY	<400>265	VQIVRQQSYS	<400>266
QIVRQQSYSI	<400>267	IVRQQSYSIM	<400>268
VRQQSYSIMS	<400>269	RQQSYSIMSI	<400>270
QQSYSIMSI	<400>271	QSYSIMSIK	<400>272
SYSIMSIKE	<400>273	YSIMSIKEE	<400>274
SIMSIKEEV	<400>275	IMSIKEEVL	<400>276
MSIKEEVL	<400>277	SIIKEEVLAY	<400>278
IIKEEVLAYV	<400>279	IKEEVLAYVV	<400>280
KEEVLAYVVQ	<400>281	EEVLAYVVQL	<400>282
EVLAYVVQLP	<400>283	VLAYVVQLPL	<400>284
LAYVVQLPLY	<400>285	AYVVQLPLYG	<400>286
YVVQLPLYGV	<400>287	VVQLPLYGVI	<400>288
VQLPLYGVID	<400>289	QLPLYGVIDT	<400>290
LPLYGVIDTP	<400>291	PLYGVIDTPC	<400>292
LYGVIDTPCW	<400>293	YGVIDTPCWK	<400>294
GVIDTPCWKL	<400>295	VIDTPCWKLH	<400>296
IDTPCWKLHT	<400>297	DTPCWKLHTS	<400>298
TPCWKLHTSP	<400>299	PCWKLHTSPL	<400>300
CWKLHTSPLC	<400>301	WKLHTSPLCT	<400>302
KLHTSPLCTT	<400>303	LHTSPLCTTN	<400>304
HTSPLCTTNT	<400>305	TSPLCTTNTK	<400>306
SPLCTTNTKE	<400>307	PLCTTNTKEG	<400>308
LCTTNTKEGS	<400>309	CTTNTKEGSN	<400>310
TTNTKEGSNI	<400>311	TNTKEGSNIC	<400>312
NTKEGSNICL	<400>313	TKEGSNICLT	<400>314
KEGSNICLTR	<400>315	EGSNICLTRT	<400>316
GSNICLTRTD	<400>317	SNICLTRTDR	<400>318
NICLTRTDRG	<400>319	ICLTRTDRGW	<400>320
CLTRTDRGWY	<400>321	LTRTDRGWYC	<400>322
TRTDRGWYCD	<400>323	RTDRGWYCDN	<400>324
TDRGWYCDNA	<400>325	DRGWYCDNAG	<400>326
RGWYCDNAGS	<400>327	GWYCDNAGSV	<400>328
WYCDNAGSVS	<400>329	YCDNAGSVSF	<400>330
CDNAGSVSFF	<400>331	DNAGSVSFFP	<400>332
NAGSVSFFPQ	<400>333	AGSVSFFPQA	<400>334
GSVSFFPQAE	<400>335	SVSFFPQAET	<400>336
VSFFPQAETC	<400>337	SFFPQAETCK	<400>338
FFPQAETCKV	<400>339	FPQAETCKVQ	<400>340
PQAETCKVQS	<400>341	QAETCKVQSN	<400>342

- 59 -

AETCKVQSNR	<400>343	ETCKVQSNRV	<400>344
TCKVQSNRVF	<400>345	CKVQSNRVFC	<400>346
KVQSNRVFCD	<400>347	VQSNRVFCDT	<400>348
QSNRVFCDTM	<400>349	SNRVFCDTMN	<400>350
NRVFCDTMNS	<400>351	RVFCDTMNSL	<400>352
VFCDTMNSLT	<400>353	FCDTMNSLTL	<400>354
CDTMNSLTLP	<400>355	DTMNSLTLP	<400>356
TMNSLTLPSE	<400>357	MNSLTLPSEV	<400>358
NSLTLPSEVN	<400>359	SLTLPSEVNL	<400>360
LTLPSEVNLC	<400>361	TLPSEVNLCN	<400>362
LPSEVNLCNV	<400>363	PSEVNLCNVD	<400>364
SEVNLCNVDI	<400>365	EVNLCNVDIF	<400>366
VNLCNVDIFN	<400>367	NLCNVDIFNP	<400>368
LCNVDIFNPK	<400>369	CNVDIFNPKY	<400>370
NVDIFNPKYD	<400>371	VDIFNPKYDC	<400>372
DIFNPKYDCK	<400>373	IFNPKYDCKI	<400>374
FNPKYDCKIM	<400>375	NPKYDCKIMT	<400>376
PKYDCKIMTS	<400>377	KYDCKIMTSK	<400>378
YDCKIMTSKT	<400>379	DKCKIMTSKT	<400>380
CKIMTSKTDV	<400>381	KIMTSKTDVS	<400>382
IMTSKTDVSS	<400>383	MTSKTDVSSS	<400>384
TSKTDVSSSV	<400>385	SKTDVSSSVI	<400>386
KTDVSSSVIT	<400>387	TDVSSSVITS	<400>388
DVSSSVITSL	<400>389	VSSSVITSLG	<400>390
SSSVITSLGA	<400>391	SSVITSLGAI	<400>392
SVITSLGAIV	<400>393	VITSLGAIVS	<400>394
ITSLGAIVSC	<400>395	TSLGAIVSCY	<400>396
SLGAIVSCYG	<400>397	LGAIVSCYCK	<400>398
GAIVSCYGKT	<400>399	AIVSCYGKTK	<400>400
IVSCYGKTKC	<400>401	VSCYGKTKCT	<400>402
SCYGKTKCTA	<400>403	CYGKTKCTAS	<400>404
YGKTKCTASN	<400>405	GKTKCTASNK	<400>406
KTKCTASNKN	<400>407	TKCTASNKNR	<400>408
KCTASNKNRG	<400>409	CTASNKNRGI	<400>410
TASNKNRGI	<400>411	ASNKNRGI	<400>412
SNKNRGI	<400>413	NKNRGI	<400>414
KNRGI	<400>415	NRGI	<400>416
RGIIKTFSNG	<400>417	GIIKTFSNGC	<400>418
IIKTFSNGCD	<400>419	IKTFSNGCDY	<400>420
KTFSNGCDYV	<400>421	TFSNGCDYVS	<400>422
FSNGCDYVSN	<400>423	SNGCDYVSNK	<400>424
NGCDYVSNKG	<400>425	GCDYVSNKGV	<400>426
CDYVSNKGVD	<400>427	DYVSNKGVD	<400>428
YVSNKGVDTV	<400>429	VSNKGVDTVS	<400>430
SNKGVDTVSV	<400>431	NKGVDTVSVG	<400>432
KGVDTVSVGN	<400>433	GVDTVSVGNT	<400>434
VDTVSVGNTL	<400>435	DTVSVGNTLY	<400>436
TVSVGNTLYY	<400>437	VSVGNTLYYV	<400>438
SVGNTLYYVN	<400>439	VGNTLYYVVK	<400>440
GNTLYYVVKQ	<400>441	NTLYYVVKQE	<400>442
TLYYVVKQEG	<400>443	LYYVVKQEGK	<400>444
YYVVKQEGKS	<400>445	YVVKQEGKSL	<400>446
VVKQEGKSLY	<400>447	NKQEGKSLYV	<400>448
KQEGKSLYVK	<400>449	QEGKSLYVKG	<400>450

- 60 -

EGKSLYVKGE	<400>451	GKSLYVKGEP	<400>452
KSLYVKGEPI	<400>453	SLYVKGEPII	<400>454
LYVKGEPIIN	<400>455	YVKGEPIINF	<400>456
VKGEPIINFY	<400>457	KGEPIINFYD	<400>458
GEPIINFYDP	<400>459	EPIINFYDPL	<400>460
PIINFYDPLV	<400>461	IINFYDPLVF	<400>462
INFYDPLVFP	<400>463	NFYDPLVFPS	<400>464
FYDPLVFPSD	<400>465	YDPLVFPSDE	<400>466
DPLVFPSDEF	<400>467	PLVFPSDEFD	<400>468
LVFPSDEFDA	<400>469	VFPSDEFDAS	<400>470
FPSDEFDASI	<400>471	PSDEFDASIS	<400>472
SDEFDASISQ	<400>473	DEFDASISQV	<400>474
EFDASISQVN	<400>475	FDASISQVNE	<400>476
DASISQVNEK	<400>477	ASISQVNEKI	<400>478
SISQVNEKIN	<400>479	ISQVNEKINQ	<400>480
SQVNEKINQS	<400>481	QVNEKINQSL	<400>482
VNEKINQSLA	<400>483	NEKINQSLAF	<400>484
EKINQSLAFI	<400>485	KINQSLAFIR	<400>486
INQSLAFIRK	<400>487	NQSLAFIRKS	<400>488
QSLAFIRKSD	<400>489	SLAFIRKSDE	<400>490
LAFIRKSDEL	<400>491	AFIRKSDELL	<400>492
FIRKSDELLH	<400>493	IRKSDELLHN	<400>494
RKSDELLHNV	<400>495	KSDELLHNVN	<400>496
SDELLHNVNA	<400>497	DELLHNVNAG	<400>498
ELLHNVNAGK	<400>499	LLHNVNAGKS	<400>500
LHNVNAGKST	<400>501	HNVNAGKSTT	<400>502
NVNAGKSTTN	<400>503	VNAGKSTTNI	<400>504
NAGKSTTNIM	<400>505	AGKSTTNIMI	<400>506
GKSTTNIMIT	<400>507	KSTTNIMITT	<400>508
STTNIMITTI	<400>509	TTNIMITTII	<400>510
TNIMITTIII	<400>511	NIMITTIIIV	<400>512
IMITTIIIVI	<400>513	MITTIIIVII	<400>514
ITTTIIIVII	<400>515	TTTIIIVIIVI	<400>516
TIIIVIIIVIL	<400>517	IIIIVIIIVILL	<400>518
IIVIIIVILLS	<400>519	IVIIIVILLSL	<400>520
VIIIVILLSLI	<400>521	IIVILLSLIA	<400>522
IVILLSLIAV	<400>523	VILLSLIAVG	<400>524
ILLSLIAVGL	<400>525	LLSLIAVGLL	<400>526
LSLIAVGLLL	<400>527	SLIAVGLLLY	<400>528
LIAVGLLLYC	<400>529	IAVGLLLYCK	<400>530
AVGLLLYCKA	<400>531	VGLLLYCKAR	<400>532
GLLLYCKARS	<400>533	LLLYCKARST	<400>534
LLYCKARSTP	<400>535	LYCKARSTPV	<400>536
YCKARSTPVT	<400>537	CKARSTPVTL	<400>538
KARSTPVTLS	<400>539	ARSTPVTLSK	<400>540
RSTPVTLSKD	<400>541	STPVTLSKDQ	<400>542
TPVTLSKDQL	<400>543	PVTLSKDQLS	<400>544
VTLSKDQLSG	<400>545	TLSKDQLSGI	<400>546
LSKDQLSGIN	<400>547	SKDQLSGINN	<400>548
KDQLSGINNI	<400>549	DQLSGINNIA	<400>550
QLSGINNIAF	<400>551	LSGINNIAFS	<400>552
SGINNIAFSN	<400>553		

- 61 -

Even more preferably said antagonist interacts with a sequence selected from <400>88, <400>89, <400>90, <400>91, <400>92, <400>93 or <400>94.

- 5 Reference to "interacts" should be understood as a reference to any form of interaction including, but not limited to covalent bonds, hydrogen bonds, ionic bonds, van der Waals forces or any other interactive bonding mechanism.

Still without limiting the present invention to any one theory or mode of action the
10 inventors have determined that inhibition or other form of interference with cleavage at the newly identified cleavage site disclosed herein interferes with F protein functioning. Further, it is thought that the intervening sequence exhibits relevance in relation to immune recognition. Specifically, it is thought that F proteins engineered to either retain the intervening sequence or which are engineered such that the intervening sequence is
15 removed altogether exhibit altered but improved immunogenicity. Although not wishing to be constrained by theory, it is thought that in the normal physiological setting, the intervening sequence which is excised following formation of the fully functional F glycoprotein serves as an immune decoy thereby obstructing or otherwise inhibiting the induction of an immune response against the fully functional F protein.

20

Accordingly, mutating the cleavage sites comprising the F protein (at either the amino acid or encoding nucleic acid level) provides a useful tool for producing molecules which are engineered to retain the intervening sequence and which cannot undergo the normal cleavage event in order to generate the fully functional F protein. These molecules are
25 useful in a range of applications including, but not limited to, as an immunogen for use in a vaccination protocol. In addition to producing a F protein variant which cannot be cleaved, identification by the inventors of the second cleavage site now enables the synthesis of F protein molecules which lack the intervening sequence as herein defined. This is particularly useful since it is thought that the F protein which lacks the intervening
30 sequence, but which intervening sequence was not released into the circulation of the subject, will exhibit better immunogenicity than the naturally occurring F protein.

- 62 -

Accordingly, in another aspect there is provided a viral F protein variant comprising a mutation in the intervening peptide sequence wherein said variant exhibits modulated functional activity relative to wild type F protein or a derivative, homologue, analogue, chemical equivalent or mimetic of said variant.

5

More particularly, there is provided a viral F protein variant comprising a mutation in the intervening peptide sequence wherein said variant exhibits down-regulated functional activity relative to wild type F protein or a derivative, homologue, analogue, chemical equivalent or mimetic of said variant.

10

Reference to "intervening peptide sequence" should be understood to have the same meaning as hereinbefore defined.

Reference to "wild type" F protein is a reference to the forms of F protein which are
15 predominantly expressed by negative sense single stranded RNA viruses. This should be understood to include reference to the uncleaved form of the F protein, the functional activity of which includes the capacity to undergo cleavage and excision of the intervening sequence, and the fully functional F protein in respect of which the intervening sequence has been excised. It should be understood that to the extent that the subject variant
20 molecule comprises all or part of the intervening sequence, modulation of its functional activity should be assessed relative to the wild type F protein which still comprises the intervening sequence. Conversely, a variant F protein which does not comprise the intervening sequence should be assessed relative to the cleaved wild type F protein. In this regard, reference to "functional activity" should be understood as a reference to any one or
25 more of the functional activities which the subject F protein can perform including, but not limited to, its capacity to undergo cleavage or its capacity to induce an immune response.

Reference to "mutation" should be understood as a reference to any change, alteration or other modification, whether occurring naturally or non-naturally, which results in the
30 subject F protein exhibiting functional activity which is modulated relative to that of the corresponding wild type F protein.

- 63 -

The change, alteration or other modification may take any form including, but not limited to, a structural modification (such as an alteration secondary, tertiary or quaternary structure of the F protein molecule), a molecular modification (such as an addition substitutional deletion of one or more amino acids from the F protein) or a chemical
5 modification. The subject modification should also be understood to extend to the fusion, linking or binding of a proteinaceous or non-proteinaceous molecule to the F protein or to the nucleic acid molecule encoding the F protein thereby rendering the expression product functionally distinctive over the corresponding wild type F protein. It should also be understood that although it is necessary that the subject mutation is expressed by the F
10 protein expression product, the creation of the mutation may be achieved by any suitable means including either mutating a wild type F protein, synthesising a F protein variant or modifying a nucleic acid molecule encoding a wild type F protein such that the expression product of said mutated nucleic acid molecule is a F protein variant. Preferably, said mutation is a single or multiple amino acid sequence substitution, addition and/or deletion.
15 In this regard, in one preferred embodiment the subject mutation is deletion of all or part of the intervening sequence. In another preferred embodiment, the subject mutation is an amino acid substitution which renders the newly identified cleavage site inactive. By inactive is meant that the cleavage site cannot be cleaved by the enzymatic processes which normally function to activate an F protein *in vivo*.
20 In a preferred embodiment the viral F protein is a Paramyxoviridae F protein and still more preferably the subject viral F protein is of the Genus Pneumovirus and still more preferably respiratory syncytial virus.
25 In a most preferred embodiment there is provided a respiratory syncytial virus F protein variant comprising a mutation in the cleavage site defined by amino acids RARR (<400>564) wherein said variant exhibits down-regulated functional activity relative to wild type F protein or a derivative, homologue, analogue, chemical equivalent or mimetic of said variant.

30

- 64 -

Preferably said mutation comprises one or more of the amino acid substitutions selected from the following list:

- (i) R106G
- 5 (ii) A107Q
- (iii) R108G

Still more preferably said F protein variant comprises the sequence substantially as set forth in <400>565.

10

In another preferred embodiment there is provided a respiratory syncytial virus F protein variant comprising a multiple amino acid deletion from the intervening peptide sequence wherein said variant exhibits down-regulated functional activity relative to wild type F protein or a derivative, homologue, analogue, chemical equivalent of said variant.

15

It is more preferably provided that said amino acid deletion is a partial deletion of the intervening peptide sequence and more preferably a deletion of the peptide sequence

RARRELPRFMNYTLNNAKKTNTLS <400>569.

20

Still more preferably said variant comprises the amino acid sequence substantially as set forth in <400>567.

To the extent that the present invention relates to F protein variants comprising one or
25 more amino acid additions, substitutions and/or deletions, it should also be understood to extend to nucleic acid molecules encoding said variants.

Accordingly, another aspect of the present invention is directed to an isolated nucleic acid molecule selected from the list consisting of:

- 30 (i) An isolated nucleic acid molecule or derivative or equivalent thereof comprising a nucleotide sequence encoding or complementary to a sequence encoding a viral F

- 65 -

protein variant or derivative, homologue, analogue, chemical equivalent or mimetic of said variant, which variant comprises a mutation in the intervening peptide sequence wherein said variant exhibits modulated functional activity relative to wild-type F protein.

5

- (ii) An isolated nucleic acid molecule or derivative or equivalent thereof comprising a nucleotide sequence encoding or complementary to a sequence encoding a viral F protein variant or derivative, homologue, analogue, chemical equivalent or mimetic of said variant, which variant comprises a mutation in the intervening peptide sequence wherein said variant exhibits down-regulated functional activity relative to wild-type F protein.

10

- (iii) An isolated nucleic acid molecule or derivative or equivalent thereof comprising a nucleotide sequence encoding or complementary to a sequence encoding a respiratory syncytial virus F protein or derivative, homologue, analogue, chemical equivalent or mimetic of said variant, which variant comprises a mutation in the cleavage site defined by amino acids RARR wherein said variant exhibits down-regulated functional activity relative to wild-type F protein.

15

- (iv) An isolated nucleic acid molecule or derivative or equivalent thereof comprising a nucleotide sequence encoding or complementary to a sequence encoding a respiratory syncytial virus F protein variant or derivative, homologue, analogue, chemical equivalent or mimetic of said variant, which variant comprises one or more of the amino acid substitutions selected from the following list:

25

- (a) R106G
- (b) A107Q
- (c) R108G

- (v) An isolated nucleic acid molecule or derivative or analogue thereof comprising a nucleotide sequence encoding or complementary to a sequence encoding a viral F

30

- 66 -

protein variant or derivative, homologue, analogue, chemical equivalent or mimetic of said variant, which variant comprises a multiple amino acid deletion from the intervening peptide sequence wherein said variant exhibits down-regulated functional activity relative to wild-type F protein.

5

- (vi) An isolated nucleic acid molecule or derivative or analogue thereof comprising a nucleotide sequence encoding or complementary to a sequence encoding a viral F protein variant or derivative, homologue, analogue, chemical equivalent or mimetic of said variant, which variant comprises a partial deletion of the intervening peptide sequence and more preferably a deletion of the peptide sequence

10

RARRELPRFMNYTLNNAKKTNVTLS <400>569.

15

- (vii) An isolated nucleic acid molecule or derivative or analogue thereof comprising a nucleotide sequence encoding or complementary to a sequence encoding a viral F protein variant or derivative, homologue, analogue, chemical equivalent or mimetic of said variant, which variant comprises the amino acid sequence substantially as set forth in <400>567.

20

- (viii) An isolated nucleic acid molecule or derivative or analogue thereof comprising a nucleotide sequence encoding or complementary to a sequence encoding a viral F protein variant or derivative, homologue, analogue, chemical equivalent or mimetic of said variant, which variant comprises the amino acid sequence substantially as set forth in <400>565.

25

- (ix) An isolated nucleic acid molecule or derivative or analogue thereof comprising the nucleotide substantially as set forth in <400>568.

30

- (x) An isolated nucleic acid molecule or derivative or analogue thereof comprising the nucleotide substantially as set forth in <400>566.

- 67 -

In a preferred embodiment the viral F protein is a Paramyxoviridae F protein and still more preferably the subject viral F protein is of the Genus Pneumovirus and still more preferably respiratory syncytial virus.

5 The nucleic acid molecule of the subject invention may be ligated to an expression vector capable of expression in a prokaryotic cell (eg. *E.Coli*) or a eukaryotic cell (eg. yeast cells, fungal cells, insect cells, mammalian cells or plant cells). The nucleic acid molecule may be ligated or fused or otherwise associated with a nucleic acid molecule encoding another entity such as, for example, a signal peptide. It may also comprise additional nucleotide
10 sequence information fused, linked or otherwise associated with it either at the 3' or 5' terminal portions or at both the 3' and 5' terminal portions. The nucleic acid molecule may also be part of a vector, such as an expression vector. The latter embodiment facilitates production of recombinant forms of the variant F protein encompassed by the present invention.

15

The variant F protein molecule of the present invention may be derived from natural or recombinant sources or may be chemically synthesised. Methods for producing these molecules would be well known to those skilled in the art.

20 As hereinbefore provided, "derivatives" include fragments, parts, portions, variants and mimetics from natural, synthetic or recombinant sources including fusion proteins. Parts or fragments include, for example, active regions of F protein. Derivatives may be derived from insertion, deletion or substitution of amino acids. Amino acid insertional derivatives include amino and/or carboxylic terminal fusions as well as intrasequence insertions of
25 single or multiple amino acids. Insertional amino acid sequence variants are those in which one or more amino acid residues are introduced into a predetermined site in the protein although random insertion is also possible with suitable screening of the resulting product. Deletional variants are characterised by the removal of one or more amino acids from the sequence. Substitutional amino acid variants are those in which at least one
30 residue in the sequence has been removed and a different residue inserted in its place. An example of substitutional amino acid variants are conservative amino acid substitutions.

- 68 -

Conservative amino acid substitutions typically include substitutions within the following groups: glycine and alanine; valine, isoleucine and leucine; aspartic acid and glutamic acid; asparagine and glutamine; serine and threonine; lysine and arginine; and phenylalanine and tyrosine. Additions to amino acid sequences include fusions with other peptides,
5 polypeptides or proteins.

Reference to "homologues" should be understood as a reference to F protein nucleic acid molecules or proteins derived from viral strains other than the species of origin.

10 Chemical and functional equivalents of F protein nucleic acid or protein molecules should be understood as molecules exhibiting any one or more of the functional activities of these molecules and may be derived from any source such as being chemically synthesized or identified via screening processes such as natural product screening.

15 The derivatives include fragments having particular epitopes or parts of the entire protein fused to peptides, polypeptides or other proteinaceous or non-proteinaceous molecules.

Analogues contemplated herein include, but are not limited to, modification to side chains, incorporating of unnatural amino acids and/or their derivatives during peptide, polypeptide
20 or protein synthesis and the use of crosslinkers and other methods which impose conformational constraints on the proteinaceous molecules or their analogues.

Derivatives of nucleic acid sequences may similarly be derived from single or multiple nucleotide substitutions, deletions and/or additions including fusion with other nucleic acid
25 molecules. The derivatives of the nucleic acid molecules of the present invention include oligonucleotides, PCR primers, antisense molecules, molecules suitable for use in cosuppression and fusion of nucleic acid molecules. Derivatives of nucleic acid sequences also include degenerate variants.

30 Examples of side chain modifications contemplated by the present invention include modifications of amino groups such as by reductive alkylation by reaction with an aldehyde followed by reduction with NaBH_4 ; amidination with methylacetimidate; acylation with acetic anhydride; carbamoylation of amino groups with cyanate;

- 69 -

trinitrobenzylation of amino groups with 2, 4, 6-trinitrobenzene sulphonic acid (TNBS); acylation of amino groups with succinic anhydride and tetrahydrophthalic anhydride; and pyridoxylation of lysine with pyridoxal-5-phosphate followed by reduction with NaBH_4 .

- 5 The guanidine group of arginine residues may be modified by the formation of heterocyclic condensation products with reagents such as 2,3-butanedione, phenylglyoxal and glyoxal.

- The carboxyl group may be modified by carbodiimide activation *via* O-acylisourea
10 formation followed by subsequent derivitisation, for example, to a corresponding amide.

- Sulphydryl groups may be modified by methods such as carboxymethylation with iodoacetic acid or iodoacetamide; performic acid oxidation to cysteic acid; formation of a mixed disulphides with other thiol compounds; reaction with maleimide, maleic anhydride
15 or other substituted maleimide; formation of mercurial derivatives using 4-chloromercuribenzoate, 4-chloromercuriphenylsulphonic acid, phenylmercury chloride, 2-chloromercuri-4-nitrophenol and other mercurials; carbamoylation with cyanate at alkaline pH.

- 20 Tryptophan residues may be modified by, for example, oxidation with N-bromosuccinimide or alkylation of the indole ring with 2-hydroxy-5-nitrobenzyl bromide or sulphenyl halides. Tyrosine residues on the other hand, may be altered by nitration with tetranitromethane to form a 3-nitrotyrosine derivative.

- 25 Modification of the imidazole ring of a histidine residue may be accomplished by alkylation with iodoacetic acid derivatives or N-carboethoxylation with diethylpyrocarbonate.

- Examples of incorporating unnatural amino acids and derivatives during protein synthesis
30 include, but are not limited to, use of norleucine, 4-amino butyric acid, 4-amino-3-hydroxy-5-phenylpentanoic acid, 6-aminohexanoic acid, t-butylglycine, norvaline, phenylglycine, ornithine, sarcosine, 4-amino-3-hydroxy-6-methylheptanoic acid, 2-thienyl

- 70 -

alanine and/or D-isomers of amino acids. A list of unnatural amino acids contemplated herein is shown in Table 3.

TABLE 3

	Non-conventional amino acid	Code	Non-conventional amino acid	Code
5	α -aminobutyric acid	Abu	L-N-methylalanine	Nmala
	α -amino- α -methylbutyrate	Mgab	L-N-methylarginine	Nmarg
	aminocyclopropane- carboxylate	Cpro	L-N-methylasparagine	Nmasn
			L-N-methylaspartic acid	Nmasp
10	aminoisobutyric acid	Aib	L-N-methylcysteine	Nmcys
	aminonorbornyl- carboxylate	Norb	L-N-methylglutamine	Nmgln
			L-N-methylglutamic acid	Nmglu
	cyclohexylalanine	Chexa	L-N-methylhistidine	Nmhis
	cyclopentylalanine	Cpen	L-N-methylisoleucine	Nmile
15	D-alanine	Dal	L-N-methylleucine	Nmleu
	D-arginine	Darg	L-N-methyllysine	Nmlys
	D-aspartic acid	Dasp	L-N-methylmethionine	Nmmet
	D-cysteine	Dcys	L-N-methylnorleucine	Nmnle
	D-glutamine	Dgln	L-N-methylnorvaline	Nmnva
20	D-glutamic acid	Dglu	L-N-methylornithine	Nmorn
	D-histidine	Dhis	L-N-methylphenylalanine	Nmphe
	D-isoleucine	Dile	L-N-methylproline	Nmpro
	D-leucine	Dleu	L-N-methylserine	Nmser
	D-lysine	Dlys	L-N-methylthreonine	Nmthr
25	D-methionine	Dmet	L-N-methyltryptophan	Nmtrp
	D-ornithine	Dorn	L-N-methyltyrosine	Nmtyr
	D-phenylalanine	Dphe	L-N-methylvaline	Nmval
	D-proline	Dpro	L-N-methylethylglycine	Nmetg
	D-serine	Dser	L-N-methyl-t-butylglycine	Nmtbug
30	D-threonine	Dthr	L-norleucine	Nle
	D-tryptophan	Dtrp	L-norvaline	Nva
	D-tyrosine	Dtyr	α -methyl-aminoisobutyrate	Maib
	D-valine	Dval	α -methyl- -aminobutyrate	Mgab

- 72 -

	D- α -methylalanine	Dmala	α -methylcyclohexylalanine	Mchexa
	D- α -methylarginine	Dmarg	α -methylcyclopentylalanine	Mcpen
	D- α -methylasparagine	Dmasn	α -methyl- α -naphthylalanine	Manap
	D- α -methylaspartate	Dmasp	α -methylpenicillamine	Mpen
5	D- α -methylcysteine	Dmcys	N-(4-aminobutyl)glycine	Nglu
	D- α -methylglutamine	Dmgln	N-(2-aminoethyl)glycine	Naeg
	D- α -methylhistidine	Dmhis	N-(3-aminopropyl)glycine	Norn
	D- α -methylisoleucine	Dmile	N-amino- α -methylbutyrate	Nmaabu
	D- α -methylleucine	Dmleu	α -naphthylalanine	Anap
10	D- α -methyllysine	Dmlys	N-benzylglycine	Nphe
	D- α -methylmethionine	Dmmet	N-(2-carbamylethyl)glycine	Ngln
	D- α -methylornithine	Dmorn	N-(carbamylmethyl)glycine	Nasn
	D- α -methylphenylalanine	Dmphe	N-(2-carboxyethyl)glycine	Nglu
	D- α -methylproline	Dmpro	N-(carboxymethyl)glycine	Nasp
15	D- α -methylserine	Dmser	N-cyclobutylglycine	Ncbut
	D- α -methylthreonine	Dmthr	N-cycloheptylglycine	Nchep
	D- α -methyltryptophan	Dmtrp	N-cyclohexylglycine	Nchex
	D- α -methyltyrosine	Dmtty	N-cyclodecylglycine	Ncdec
	D- α -methylvaline	Dmval	N-cylcododecylglycine	Ncdod
20	D-N-methylalanine	Dnmala	N-cyclooctylglycine	Ncoct
	D-N-methylarginine	Dnmarg	N-cyclopropylglycine	Ncpro
	D-N-methylasparagine	Dnmasn	N-cycloundecylglycine	Ncund
	D-N-methylaspartate	Dnmasp	N-(2,2-diphenylethyl)glycine	Nbhm
	D-N-methylcysteine	Dnmcys	N-(3,3-diphenylpropyl)glycine	Nbhe
25	D-N-methylglutamine	Dnmgln	N-(3-guanidinopropyl)glycine	Narg
	D-N-methylglutamate	Dnmglu	N-(1-hydroxyethyl)glycine	Nthr
	D-N-methylhistidine	Dnmhis	N-(hydroxyethyl)glycine	Nser
	D-N-methylisoleucine	Dnmile	N-(imidazolylethyl)glycine	Nhis
	D-N-methylleucine	Dnmleu	N-(3-indolylyethyl)glycine	Nhtrp
30	D-N-methyllysine	Dnmlys	N-methyl- γ -aminobutyrate	Nmgabu
	N-methylcyclohexylalanine	Nmchexa	D-N-methylmethionine	Dnmmet
	D-N-methylornithine	Dnmorn	N-methylcyclopentylalanine	Nmcpen
	N-methylglycine	Nala	D-N-methylphenylalanine	Dnmphe

- 73 -

	N-methylaminoisobutyrate	Nmaib	D-N-methylproline	Dnmpro
	N-(1-methylpropyl)glycine	Nile	D-N-methylserine	Dnmser
	N-(2-methylpropyl)glycine	Nleu	D-N-methylthreonine	Dnmthr
	D-N-methyltryptophan	Dnmtrp	N-(1-methylethyl)glycine	Nval
5	D-N-methyltyrosine	Dnmtyr	N-methyl- α -naphthylalanine	Nmanap
	D-N-methylvaline	Dnmval	N-methylpenicillamine	Nmpen
	γ -aminobutyric acid	Gabu	N-(<i>p</i> -hydroxyphenyl)glycine	Nhtyr
	L- <i>t</i> -butylglycine	Tbug	N-(thiomethyl)glycine	Ncys
	L-ethylglycine	Etg	penicillamine	Pen
10	L-homophenylalanine	Hphe	L- α -methylalanine	Mala
	L- α -methylarginine	Marg	L- α -methylassparagine	Masn
	L- α -methylaspartate	Masp	L- α -methyl- <i>t</i> -butylglycine	Mtbug
	L- α -methylcysteine	Mcys	L-methylethylglycine	Metg
	L- α -methylglutamine	Mgln	L- α -methylglutamate	Mglu
15	L- α -methylhistidine	Mhis	L- α -methylhomophenylalanine	Mhphe
	L- α -methylisoleucine	Mile	N-(2-methylthioethyl)glycine	Nmet
	L- α -methylleucine	Mleu	L- α -methyllysine	Mlys
	L- α -methylmethionine	Mmet	L- α -methylnorleucine	Mnle
	L- α -methylnorvaline	Mnva	L- α -methylornithine	Morn
20	L- α -methylphenylalanine	Mphe	L- α -methylproline	Mpro
	L- α -methylserine	Mser	L- α -methylthreonine	Mthr
	L- α -methyltryptophan	Mtrp	L- α -methyltyrosine	Mtyr
	L- α -methylvaline	Mval	L-N-methylhomophenylalanine	Nmhph
	N-(N-(2,2-diphenylethyl)	Nnbhm	N-(N-(3,3-diphenylpropyl)	Nnbhe
25	carbamylmethyl)glycine		carbamylmethyl)glycine	
	1-carboxy-1-(2,2-diphenyl-Nmbc			
	ethylamino)cyclopropane			

-
- 30 Crosslinkers can be used, for example, to stabilise 3D conformations, using homo-bifunctional crosslinkers such as the bifunctional imido esters having (CH₂)_n spacer groups with n=1 to n=6, glutaraldehyde, N-hydroxysuccinimide esters and hetero-bifunctional

- 74 -

reagents which usually contain an amino-reactive moiety such as N-hydroxysuccinimide and another group specific-reactive moiety.

In addition to screening for agents which modulate F protein functional activity, the development of a method of producing a viral F protein or derivative thereof in a eukaryotic cell and identification of the novel F protein cleavage site has now facilitated the development of *in vivo* methodology directed to administering to a subject a vaccine comprising a nucleic acid molecule encoding a viral F protein or derivative thereof. Reference to "derivative" should be understood to encompass variants thereof, such as the variants hereinbefore defined. Without limiting the present invention to any one theory or mode of action, the operation of such a vaccine is based on the generation of an immune response, in particular antibody synthesis, directed to the subject F protein or derivative thereof. The antibodies generated therein bind to virally produced F proteins thereby inhibiting their fusion related functional activity and consequently reducing and/or inhibiting further viral propagation. Such a vaccine is useful in either the prophylactic and/or therapeutic sense.

Accordingly, another aspect of the present invention provides a recombinant viral construct comprising a nucleic acid molecule encoding a viral F protein or derivative thereof, the nucleotide sequence of which nucleic acid molecule comprises codons optimised for expression in a eukaryotic cell, wherein said recombinant viral construct is effective in inducing, enhancing or otherwise stimulating an immune response to said F protein.

Still another aspect of the present invention provides a recombinant viral construct comprising a nucleic acid molecule encoding a viral F protein variant or derivative thereof wherein said recombinant viral construct is effective in inducing, enhancing or otherwise stimulating an immune response to said F protein variant.

- 75 -

In a preferred embodiment the viral F protein is a Paramyxoviridae F protein and still more preferably the subject viral F protein is of the Genus Pneumovirus and still more preferably respiratory syncytial virus.

- 5 Reference to "inducing, enhancing or otherwise stimulating" an immune response to an F protein should be understood to mean stimulating or facilitating the stimulation of a specific immune response. The specific immune response is preferably a humoral response which is directed to any one or more regions of the F protein. In this regard, it should be understood that the subject immune response will down-regulate and/or inhibit
10 at least one functional activity of the subject F protein.

- Yet another aspect of the present invention relates to a vaccine comprising a recombinant viral construct which construct comprises a nucleic acid molecule encoding a respiratory syncytial virus F protein or derivative thereof, the nucleotide sequence of which nucleic
15 acid molecule is optimised for expression in a eukaryotic cell wherein said recombinant viral construct is effective in inducing, enhancing or otherwise stimulating an immune response to said F protein.

- Still another aspect of the present invention relates to a vaccine comprising a recombinant
20 viral construct which construct comprises a nucleic acid molecule encoding a respiratory syncytial virus F protein variant or derivative thereof, wherein said recombinant viral construct is effective in inducing, enhancing or otherwise stimulating an immune response to said F protein variant.

- 25 In accordance with these aspects of the present invention, the nucleotide sequence of the subject nucleic acid molecule is preferably the nucleotide sequence defined in <400>5, <400>6, <400>566 or <400>568.

- A further aspect of the present invention relates to use of the agents hereinbefore defined to
30 modulate F protein functional activity and, in particular, the use of these agents in the therapeutic and/or prophylactic treatment of conditions characterised by infection with a

- 76 -

negative sense single stranded RNA virus, and more particularly respiratory syncytial virus. Conditions envisaged herein include Parainfluenza induced croup and bronchiolitis. It should be understood that reference to "agent" hereinafter includes reference to agents identified or generated by the screening assays described above, including the modulatory
5 agents (for example, antibodies) which are generated *in vivo* via use of a DNA vaccine. This aspect of the present invention is also directed to use of the F protein or derivatives thereof or encoding nucleic acid molecules, including the F protein variants, as hereinbefore described in the therapeutic and/or prophylactic treatment of conditions characterised by infection with a negative sense single stranded RNA virus.

10

Accordingly, another aspect of the present invention provides the method of modulating at least one functional activity associated with a viral F protein in a subject, said method comprising introducing into said subject and effective amount of an F protein modulatory agent for a time and under condition sufficient for said agent to interact with said F protein.

15

Preferably, said functional activity is F protein mediated host cell-virion fusion and/or virion budding and said modulation is down-regulation.

In a preferred embodiment the viral F protein is a Paramyxoviridae F protein and still more
20 preferably the subject viral F protein is of the Genus Pneumovirus and still more preferably respiratory syncytial virus.

The term "subject" includes humans primates, livestock animals(eg, horses, cattle, sheep, pigs, donkeys), laboratory test animals (eg, mice, rats, rabbits, guinea pigs), companion
25 animals (eg, dogs, cats), captive wild animals (eg, kangaroos, deer, foxes), birds (eg, chickens, ducks, bantams, pheasants). Preferably the subject is a human or laboratory test animal. Even more preferably the subject is a human.

In another aspect, the present invention provides a method of modulating at least one
30 functional activity associated with a viral F protein, said method comprising contacting

- 77 -

said viral F protein with an effective amount of an F protein modulatory agent for a time and under conditions sufficient for said agent to interact with said F protein.

Preferably said viral F protein is a Pneumovirus F protein and even more preferably a
5 respiratory syncytial virus F protein. Still more preferably said modulation is down-regulation of F protein functional activity.

This aspect of the present invention should be understood to extend to the modulation of F protein associated functional activities in *in vitro* culture systems. This may be of benefit,
10 for example, when applied to *in vitro* procedures designed to virally infect a prospective host cell. This may be of particular use, for example, where it is desired to create a cell line or to otherwise create a virally transformed cell. In this regard, the subject modulation would preferably be up-regulation of F protein functional activity.

15 In yet another aspect, the present invention relates to a method for the treatment and/or prophylaxis of a condition characterised by infection with a negative sense single stranded RNA virus in a subject, said method comprising administering to said subject an effective amount of an agent, which agent is capable of down-regulating at least one functional activity of the F protein expressed by said virus, for a time and under conditions sufficient
20 for said agent to interact with said F protein.

In still yet another aspect, the present invention relates to a method for the treatment and/or prophylaxis of a condition characterised by infection with a negative sense single stranded RNA virus in a subject, said method comprising administering to said subject an effective
25 amount of a composition comprising an F protein or derivative thereof, F protein variant or derivative thereof and/or a nucleic acid molecule encoding said F protein or F protein variant as hereinbefore defined or a derivative, homologue, analogue, chemical equivalent a mimetic of said protein or nucleic acid molecule for a time and under conditions sufficient for said composition to down-regulate said viral F protein functional activity.

30

- 78 -

In a preferred embodiment the viral F protein is a Paramyxoviridae F protein and still more preferably the subject viral F protein is of the Genus Pneumovirus and still more preferably respiratory syncytial virus.

- 5 Reference to "administering" an agent should be understood to extend to the administration of a DNA vaccine for the purpose of *in vivo* generation of anti - F protein antibodies.

Reference to a condition "characterised by infection with a negative sense single stranded RNA virus" should be understood as a reference to a condition, one or more symptoms of
10 which are directly or indirectly induced due to infection of the subject with the subject virus. Preferably, said virus is a Pneumovirus and even more preferably respiratory syncytial virus.

The molecule which may be administered to a subject in accordance with the present
15 invention may also be linked to a targeting means such as a monoclonal antibody, which provides specific delivery of the molecule to the target cells.

In a preferred embodiment the subject of the prophylactic or therapeutic treatment is a mammal and still more preferably a human.

20

Administration of the subject modulatory agent or the subject F protein or derivative thereof, F protein variant or derivative thereof, nucleic acid molecule encoding said F protein or F protein variant as hereinbefore defined or a derivative, homologue, analogue, chemical equivalent or mimetic of said protein or nucleic acid molecule (hereinafter said
25 modulatory agents, proteins and/or nucleic acid molecules are collectively referred to as the "active ingredients"), in the form of a pharmaceutical composition, may be performed by any convenient means. The active ingredients of the pharmaceutical composition are contemplated to exhibit therapeutic activity when administered in an amount which depends on the particular case. The variation depends, for example, on the human or
30 animal and the active ingredient chosen. A broad range of doses may be applicable. Considering a patient, for example, from about 0.1 mg to about 1 mg of active ingredient

- 79 -

may be administered per kilogram of body weight per day. Dosage regimes may be adjusted to provide the optimum therapeutic response. For example, several divided doses may be administered daily, weekly, monthly or other suitable time intervals or the dose may be proportionally reduced as indicated by the exigencies of the situation. The active
5 ingredient may be administered in the form of pharmaceutically acceptable nontoxic salts, such as acid addition salts or metal complexes, e.g. with zinc, iron or the like (which are considered as salts for purposes of this application). Illustrative of such acid addition salts are hydrochloride, hydrobromide, sulphate, phosphate, maleate, acetate, citrate, benzoate, succinate, malate, ascorbate, tartrate and the like. If the active ingredient is to be
10 administered in tablet form, the tablet may contain a binder such as tragacanth, corn starch or gelatin; a disintegrating agent, such as alginic acid; and a lubricant, such as magnesium stearate.

Routes of administration include, but are not limited to, respiratorally, intratracheally,
15 nasopharyngeally, intravenously, intraperitoneally, subcutaneously, intracranially, intradermally, intramuscularly, intraocularly, intrathecally, intracerebrally, intranasally, infusion, orally, rectally, *via* IV drip patch and implant. Preferably, the route of administration is a route which permits directed delivery of the modulatory agent. For example, aerosol administration (such as by nebulisation) into the airways permits directed
20 delivery to the airways region, in contrast to systemic delivery which results in delivery to the whole body.

Where the disorder which is the subject of treatment or prophylaxis is a respiratory distress syndrome, delivery of the active ingredient to the airway, for example as an aerosol *via*
25 nebulisation, is an ideal approach since this maximises delivery to the airway where the infection has occurred and minimises systemic delivery which may be associated with side effects.

The term "aerosol" is used in its most general sense to include any formulation capable of
30 administration *via* nasal, pharyngeal, tracheal, bronchial or oral passages. Aerosols generally comprise particles of liquid or solid suspended in a gas or vapour. Conveniently,

- 80 -

the aerosol is a colloidal system such as a mist in which the dispersion medium is a gas. The method of administering the aerosol formulation is not critical and may be achieved using a nasal spray hand pump, electric pump, pressurised dispenser, nasal drip or other convenient means. Alternatively, the formulation may be administered in a dry powder
5 delivery system. It should be understood that the method of the present invention extends to direct application of said formulations to intra nasal surfaces. In a particularly preferred embodiment, the aerosol is delivered at a rate of from about 1 to about 20 litres/min. and preferably from about 2 to about 15 litres/min. at a droplet size of from about 0.1 to about 10 μm and more preferably from about 0.1 to about 6 μm . Conveniently, a stock solution
10 of material is prepared at a concentration of from about 0.5 to about 20 mg/ml or more preferably from about 1.0 to about 10 mg/ml of carrier solution.

The formulation is administered in a therapeutically effective amount. A therapeutically effective amount means that amount necessary at least partly to attain the desired effect, or
15 to delay the onset of, inhibit the progression of, or halt altogether, the onset or progression of the particular condition being treated. Such amounts will depend, of course, on the particular conditions being treated, the severity of the condition and individual patient parameters including age, physical conditions, size, weight and concurrent treatment. These factors are well known to those of ordinary skill in the art and can be addressed with
20 no more than routine experimentation. It is preferred generally that a maximum dose be used, that is, the highest safe dose according to sound medical judgement. It will be understood by those of ordinary skill in the art, however, that a lower dose or tolerable dose may be administered for medical reasons, psychological reasons or for virtually any other reasons.

25 Generally, daily doses of formulation will be from about 0.01 $\mu\text{g/kg}$ per day to 1000 mg/kg per day. Small doses (0.01-1 mg) may be administered initially, followed by increasing doses up to about 1000 mg/kg per day. In the event that the response in a subject is insufficient at such doses, even higher doses (or effective higher doses by a different, more
30 localised delivery route) may be employed to the extent patient tolerance permits. A single dose may be administered or multiple doses may be required on an hourly, daily, weekly or

- 81 -

monthly basis. Effective amounts of formulation vary depending on the individual but may range from about 0.1 µg to about 20 mg, alternatively from about 1 µg to about 10 mg and more preferably from about 1 µg to 5 mg per dose.

5 In another aspect the present invention relates to the use of an agent capable of modulating at least one functional activity of a viral F protein, which agent is identified and/or generated in accordance with the methods hereinbefore defined, in the manufacture of a medicament for the treatment and/or prophylaxis of a condition characterised by infection with a negative sense single stranded RNA virus.

10

In still another aspect the present invention relates to the use of a composition comprising an F protein or derivative thereof, F protein variant or derivative thereof, nucleic acid molecule encoding said F protein or F protein variant as hereinbefore defined or a derivative, homologue, analogue, chemical equivalent or mimetic of said protein or nucleic acid molecule, in the manufacture of a medicament for the treatment and/or prophylaxis of a condition characterised by infection with a negative sense single stranded RNA virus.

15

In a preferred embodiment the viral F protein is a Paramyxoviridae F protein and still more preferably the subject viral F protein is of the Genus Pneumovirus and still more preferably respiratory syncytial virus.

20

In another aspect the present invention relates to the use of an agent, which agent is identified in accordance with the methods hereinbefore defined, in the manufacture of a medicament for the modulation of at least one viral F protein associated functional activity.

25

In a preferred embodiment the viral F protein is a Paramyxoviridae F protein and still more preferably the subject viral F protein is of the Genus Pneumovirus and still more preferably respiratory syncytial virus.

- 82 -

Yet another aspect relates to agents for use in modulating the functional activity of a viral F protein wherein said agent is identified in accordance with the methods hereinbefore defined.

- 5 Still yet another aspect relates to agents for use in the treatment and/or prophylaxis of a condition characterised by infection with a negative sense single stranded RNA virus wherein said agent is identified in accordance with the methods hereinbefore defined.

- 10 Yet still another aspect relates to a composition comprising an F protein or derivative thereof, F protein variant or derivative thereof, a nucleic acid molecule encoding said F protein or F protein variant as hereinbefore defined or a derivative, homologue, analogue, chemical equivalent or mimetic of said protein or nucleic acid molecule for use in the treatment and/or prophylaxis of a condition characterised by infection with a negative sense single stranded RNA virus.

15

In a preferred embodiment the viral F protein is a Paramyxoviridae F protein and still more preferably the subject viral F protein is of the Genus Pneumovirus and still more preferably respiratory syncytial virus.

- 20 Reference herein to "treatment" and "prophylaxis" is to be considered in its broadest context. The term "treatment" does not necessarily imply that a mammal is treated until total recovery. Similarly, "prophylaxis" does not necessarily mean that the subject will not eventually contract a disease condition. Accordingly, treatment and prophylaxis include amelioration of the symptoms of a particular condition or preventing or otherwise reducing
25 the risk of developing a particular condition. The term "prophylaxis" may be considered as reducing the severity of onset of a particular condition. "Treatment" may also reduce the severity of an existing condition or the frequency of acute attacks.

- 30 In accordance with these methods, the active ingredients defined in accordance with the present invention may be coadministered with one or more other compounds or molecules. By "coadministered" is meant simultaneous administration in the same formulation or in

- 83 -

two different formulations via the same or different routes or sequential administration by the same or different routes. By "sequential" administration is meant a time difference of from seconds, minutes, hours or days between the administration of the two types of molecules. These molecules may be administered in any order.

5

In yet another aspect the present invention relates to a pharmaceutical composition comprising an active ingredient, as hereinbefore defined, and one or more pharmaceutically acceptable carriers and/or diluents.

- 10 The pharmaceutical forms suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion or may be in the form of a cream or other form suitable for topical application. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of
- 15 microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol and liquid polyethylene glycol, and the like), suitable mixtures thereof, and vegetable oils. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion
- 20 and by the use of surfactants. The preventions of the action of microorganisms can be brought about by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, thimerosal and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars or sodium chloride. Prolonged absorption of the injectable compositions can be brought about by the use in the
- 25 compositions of agents delaying absorption, for example, aluminum monostearate and gelatin.

- Sterile injectable solutions are prepared by incorporating the active compounds in the required amount in the appropriate solvent with various of the other ingredients
- 30 enumerated above, as required, followed by filtered sterilisation. Generally, dispersions are prepared by incorporating the various sterilised active ingredient into a sterile vehicle which contains the basic dispersion medium and the required other ingredients from those

- 84 -

enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum drying and the freeze-drying technique which yield a powder of the active ingredient plus any additional desired ingredient from previously sterile-filtered solution thereof.

- 5 When the active ingredients are suitably protected they may be orally administered, for example, with an inert diluent or with an assimilable edible carrier, or it may be enclosed in hard or soft shell gelatin capsule, or it may be compressed into tablets, or it may be incorporated directly with the food of the diet. For oral therapeutic administration, the
- 10 active compound may be incorporated with excipients and used in the form of ingestible tablets, buccal tablets, troches, capsules, elixirs, suspensions, syrups, wafers, and the like. Such compositions and preparations should contain at least 1% by weight of active compound. The percentage of the compositions and preparations may, of course, be varied and may conveniently be between about 5 to about 80% of the weight of the unit. The
- 15 amount of active compound in such therapeutically useful compositions is such that a suitable dosage will be obtained. Preferred compositions or preparations according to the present invention are prepared so that an oral dosage unit form contains between about 0.1 μ g and 2000 mg of active compound.
- 20 The tablets, troches, pills, capsules and the like may also contain the components as listed hereafter: a binder such as gum, acacia, corn starch or gelatin; excipients such as dicalcium phosphate; a disintegrating agent such as corn starch, potato starch, alginic acid and the like; a lubricant such as magnesium stearate; and a sweetening agent such as sucrose, lactose or saccharin may be added or a flavouring agent such as peppermint, oil of
- 25 wintergreen, or cherry flavouring. When the dosage unit form is a capsule, it may contain, in addition to materials of the above type, a liquid carrier. Various other materials may be present as coatings or to otherwise modify the physical form of the dosage unit. For instance, tablets, pills, or capsules may be coated with shellac, sugar or both. A syrup or elixir may contain the active compound, sucrose as a sweetening agent, methyl and
- 30 propylparabens as preservatives, a dye and flavouring such as cherry or orange flavour. Of course, any material used in preparing any dosage unit form should be pharmaceutically

- 85 -

pure and substantially non-toxic in the amounts employed. In addition, the active compound(s) may be incorporated into sustained-release preparations and formulations.

Pharmaceutical compositions suitable for aerosol administration have been hereinbefore
5 described.

The pharmaceutical composition may also comprise genetic molecules such as a vector capable of transfecting target cells where the vector carries a nucleic acid molecule encoding an active ingredient. The vector may, for example, be a viral vector.

10

The present invention is further described by the following non-limiting examples.

- 86 -

TABLE 1

	Sequence ID Number	Description
	<400>1	Natural F protein nucleic acid sequence
	<400>2	Natural F _{sol} portion nucleic acid sequence
5	<400>3	Restriction site modified F protein nucleic acid sequence
	<400>4	Restriction site modified F _{sol} portion nucleic acid sequence
	<400>5	Splice site and codon optimised F protein nucleic acid sequence
	<400>6	Splice site and codon optimised F _{sol} portion nucleic acid sequence
10	<400>7	F protein amino acid sequence
	<400>8	F _{sol} portion amino acid sequence
	<400>9 - <400>553	F protein amino acid decapeptides
	<400>554	P protein amino acid sequence
15	<400>555	Natural P protein nucleic acid sequence
	<400>556	Optimised P protein nucleic acid sequence
	<400>557	N protein amino acid sequence
	<400>558	Natural N protein nucleic acid sequence
	<400>559	Optimised N protein nucleic acid sequence
20	<400>560	SH protein amino acid sequence
	<400>561	Natural SH protein nucleic acid sequence
	<400>562	Optimised SH protein nucleic acid sequence
	<400>563	F protein cleavage site 1 aa sequence
	<400>564	F protein cleavage site 2 aa sequence
25	<400>565	F protein variant
	<400>566	F protein variant nucleic acid sequence
	<400>567	F protein variant
	<400>568	F protein variant nucleic acid sequence
	<400>569	F protein intervening aa sequence
30	<400>570	Poly (a) adenylation site

- 87 -

EXAMPLE 1

DESIGN OF SYNTHETIC GENE FOR RSV F EXPRESSION

Initial attempts to express the RSV F gene sequence in a soluble form (truncated at the transmembrane domain) proved unsuccessful in achieving high levels of expression. The sequence used in the expression vectors was called F_{sol}. (this differed from the viral sequence in 24/1575 nucleotides where restriction sites had been inserted to allow for easy mutagenesis – see Fig. 2b). The F viral sequence (F_{sol}.viral Fig 2b) contained suboptimal codon usage for expression in mammalian cells. In addition, a possible eight 3' splice sites were identified, including preceding lariat sequences at four positions. Poly (A) adenylation sites (AATAAA <400>570) were also identified at 4 positions. In addition, the F natural sequence like the viral sequence is approximately 65% AT rich. Most mammalian expressed genes are less than 50% AT rich. The DNA sequence encoding the transmembrane form of RSV F is also shown in Fig 2a.

In an attempt to overcome poor expression levels in mammalian cells, a new F sequence was designed that:

- (a) retained the same encoded amino acid sequence
- (b) used wherever possible optimum codon usage
- (c) removed all potential splice sites and poly A sites
- (d) removed as many CG doublets as these may be methylation sites
- (e) designed unique restriction sites to allow cassette mutagenesis
- (f) sequence was checked by secondary structure and any large hairpin loops were destabilised by changing the sequence

Sequences encoding a transmembrane version of F and the F_{sol} protein are shown in Fig.3a and 3b respectively.

Both of these optimised sequences F_{opt} and F_{sol}.opt are compared to the viral sequence in Figs 2a and 2b.

- 88 -

The synthetic DNA sequence Fopt (also referred to as F_{sol.opt}) was assembled and cloned as outlined in Fig. 4a and 4b. In brief, single stranded synthetic DNA fragments of average length 60 bases were annealed and ligated together to produce three fragments

- 5 (1) a 631bp Pst I-Mfe I fragment
- (2) a 606bp Mfe I-Xho I fragment
- (3) a 379bp Xho I-Bam HI fragment.

These gel purified fragments were cloned in pLitmus 38 or a derivative of pLITMUS
10 (pLITMUS 273/279). Clones containing the correct sequence were used as a DNA source to assemble the full length gene as outlined in Fig. 4b. In brief the respective fragment Pst-Mfe I, Xho I-Bam HI and Mfe-Xho I were sequentially cloned into the CMV expression vector pCICO or its derivatives. [pCICO is a derivative of pJW4304 which contains a full length CMV promoter and the CMV authentic intron sequence preceding the Pst I site.
15 The 3' terminator used is derived from SV40 early region and this vector also contains the SV40 origin of replication. The plasmid is from the pUC series and contains an ampicillin resistance gene. (pJW4304 was obtained from J. Mullins Dept. of Microbiology, University of Washington, Chapman *et al.*, NAR, 19:3979-3980, 1991)]. This produced the final clone pCICO.Fopt.

20

pCICO.Fopt was further modified by cloning in a 270bp EcoRI-Xba I fragment (see Fig. 4b) which encodes the transmembrane and cytoplasmic domains of the RSV F protein. Again, the DNA sequence was optimised as for the soluble version See Fig. 2b for comparison of F.opt (Fopt FL sequence) and F (viral with a few additional restriction site
25 changes) and F.viral (viral sequence). The resulting CMV expression plasmid is called pCICO.F.FL.opt. Note FL stands for the term full length and refers to a form of F that includes the transmembrane region and the cytoplasmic tail.

30

- 89 -

EXAMPLE 2***IN VITRO* EXPRESSION OF RSV F EXPRESSION**

Vectors pCICO containing the F_{sol}-opt sequence (pCICO.Fopt) and the F_{sol} sequence
5 (pCICO.FS3) were tested for expression by CaPO₄ precipitation in 293 cells. Cells in a
60ml dish were transfected with 5µg of plasmid and 0.5µg of pVARNA. Cells were
radioactively labelled with ³⁵S methionine and ³⁵S cysteine 24 hours post transfection and
the supernatants collected 5 hours after labelling. Supernatants were immunoprecipitated
with a RSV F specific monoclonal antibody and the precipitates were analysed by
10 polyacrylamide gel electrophoresis. Gels were subjected to fluorography, dried and
exposed to X-ray film. Fig. 5 shows an autoradiograph comparing the amount of F in
pCICO.FS3, pCICO.Fopt and control (mock-transfected) cells. Expression is much
improved in the pCICO.Fopt transfected cells by at least 20 fold.

15

EXAMPLE 3**RSV FUSION ASSAY**

293 cells were also transfected with the plasmid pCICO.F.FL.opt which contains the
transmembrane spanning version of F. Cells transfected with this plasmid were observed
20 24-48 hours post transfection to contain many large syncytia and dying cells. Control cells
were confluent. The F transfected cells look indistinguishable from RSV infected cells.
Thus high level expression of F is all that is necessary for cell fusion to occur. This is
markedly different to what is reported in the literature (Collins et al, Fields, and references
within). This assay forms a useful screen for detecting F specific inhibitors of RSV fusion.
25 Agents found by this assay are also useful for inhibiting RSV replication.

EXAMPLE 4**RSV SECOND CLEAVAGE SITE MUTANTS**

30 The RSV F protein sequence at amino acid singular numbers 106-109, contains the
sequence RARR. As shown in Figure 1c, this potential cleavage site is contained within

- 90 -

the F2 sub-unit of the F protein. When the F protein is expressed in mammalian cells, proteolytic cleavage occurs at two sites being site 1 (KKRKRR amino acids 131-136) which was previously identified and the previously unknown site 2 (RARR amino acids 106-109).

5

The site RARR was mutated to GQGR in the expression plasmid pCICO.FL.Fopt to give rise to the plasmid pCICO.F.FL.S2-2. Transfection of this plasmid into 293 cells revealed cleavage at site 1 but not at site 2 as expected. This was detected by a larger size F2 sub-unit (~30K versus 18K) in the S2-2 mutant than in the wild type. The size of the protein
10 between site 2 and site 1 would be expected to be 10-12K (25 amino acids plus two NH₂ - linked glycosylation sites). It was surprisingly noted that no evidence of fusion was seen in the 293 cells transfected with the S2-2 mutant plasmid of wild type. This evidence would suggest that cleavage at both site 1 and site 2 is necessary for cleavage. Note that in additional experiments, mutation of site 1 (KKRKRR) to GGKQGR, produced a mutant
15 showing no fusion activity.

In the next experiments the issue of whether the sequence between sites 1 and 2 were necessary for fusion was addressed. A mutant was constructed by standard techniques (cassette mutagenesis) in which amino acids 106-130 were deleted. This mutant is
20 designated delta 106-130. Transfection of 293 cells with an expression plasmid containing this mutant (pCICO.FLFA106-130) showed that fusion did occur. This fusion was phenotypically different from wild type in that only small syncytia were visible, suggesting that the ability of the RSV F protein to initiate or perform fusion had been attenuated.

25

- 91 -

EXAMPLE 5

EXPRESSION OF NATURAL F -V- F OPTIMISED SEQUENCE

5 *Cloning of RSVA2 F cDNA*

RNA prepared from RSVA2 infected Hep-2 cells was used as a source of RSV A2 F mRNA. RT-PCR (reverse transcriptase PCR) using 5'- and 3'- end primers was used to prepare cDNA encoding RSV A2 F according to standard methods. PCR products were
10 subcloned into standard vectors. Sequencing of many clones revealed a consensus sequence for the F gene of RSV A2. This sequence is shown in Figure 6 as F.nat and compared to F .viral. The F.nat sequence differs at nt 174 and 222. Both of these T to C changes do not result in amino acid changes. A pCICO vector containing the F.nat sequence (called pCICO.F.nat) was assembled from a synthetic Pst1 to Acc1 157 bp
15 fragment ligated to a 445 bp Acc1 to Mfe 1 fragment and a 1125 bp Mfe 1 to Xba 1 fragment derived from independent RT-PCR RSVA2 F cDNA clones. The synthetic fragment was used to make the addition of extra 5'-untranslated sequences not present in the PCR products. The 5'-untranslated sequence is 5'- CTGCAGTCACCGTCCTTGACACC -3' (<400>571) and includes a Pst 1 site. This sequence is added just 5' to the
20 initiator ATG in the following constructions pCICO.F.nat and the previously described pCICO.F.FL.opt. The Acc1 to Mfe 1 and Mfe 1 to Xba1 fragments were derived from independent RT-PCR RSVA2 F cDNA clones. The sequence F.nat encodes the same 574 amino acid sequence as shown in Fig 1.

25 *Expression of pCICO.F.FL.opt versus pCICO.F.nat*

293 cells were transfected with plasmids pCICO.F.FL.opt , pCICO.F.nat and a control as described in example 2. Cells were harvested at 24, 48 and 72 hours post transfection in cell lysis buffer. The amount of F protein in these samples was measured by Western blot
30 analysis using standard techniques. The primary antibody called 18B2, is a mouse monoclonal antibody that recognizes the F1 protein. A proteolytic breakdown product of

- 92 -

F1 called F1' is also recognized by this antibody. The western blots were developed using a secondary anti - mouse horseradish peroxidase antibody and a light emitting substrate according to standard procedures.

- 5 The results of these experiments are shown in fig 7. Lanes labelled WT refer to samples from cells transfected with pCICO.F.FL.opt : A2 lanes refer to samples from cells transfected with pCICO.F.nat and Ctrl lanes are from cells transfected with control plasmids lacking either F sequence. F protein (F1 and F1') is only observed in WT lanes indicating that the F expression level in cells transfected with pCICO.F.FL.opt is far
10 superior to those transfected with pCICO.F.nat.

- In parallel to the above experiments 293 cells were transfected with the same three plasmids and observed microscopically for signs of cell to cell fusion (syncytia formation). In three parallel experiments only cells transfected with pCICO.F.FL.opt show any cell to
15 cell fusion. At 72 hours post transfection between 75 to 100 % of cells were involved in syncytia in pCICO.FL.opt transfected cells. No fusion is observed in either the pCICO.F.nat or Ctrl transfected cells (see Fig 8).

- Those skilled in the art will appreciate that the invention described herein is susceptible to
20 variations and modifications other than those specifically described. It is to be understood that the invention includes all such variations and modifications. The invention also includes all of the steps, features, compositions and compounds referred to or indicated in this specification, individually or collectively, and any and all combinations of any two or more of said steps or features.

BIBLIOGRAPHY

Chapman, B.S., Thayer, R.M., Vincent, K.A. and Haigwood, N.L., (1991), *Nucleic Acids Res.*, **19**(14):3979-3986

Collins *et al.*, In: *Fields Virology Third Edition (Eds)*, Fields B., Knipe D., and Howley P., Lippincott Raven, Philadelphia, USA, (1996) pp 1313-1352

Gardsvoll, H., Dano, K., and Ploug, M., (1999), *J Biol Chem*, **274**(53):37995-38003

Goodson, R.J., Doyle, M.V., Kaufman, S.E., Rosenberg, S., (1994), *Proc. Natl. Acad. Sci. USA*, **91**(15):7129-7133

Haas, J., Parke E.C. and Seed B., (1996), *Current Biology*, **6**:315-324

Hoyer-Hansen, G., Behrendt, N., Ploug, M., Dano, K., and Preissner, K.T., (1997), *FEBS Lett*, **420**(1):79-85

Kuhnle *et al.*, (1998), *Journal of Virology*, **72**(5):3804-3811

Lathe, R., (1985), *J. Mol. Biol.*, **183**:1-12

Lopez *et al.*, (1998), *J. Virol.*, **72**:6922-6928

Navaza, J., (1994), *Acta Cryst*, **A50**:157-163

Ploug, M., (1998), *Biochemistry*, **37**(47):16494-16505

Ploug, M., Ellis, V., and Dano, K., (1994), *Biochemistry*, **33**(30):8991-8997

- 94 -

Ploug, M., Ostergaard, S., Hansen, L.B., Holm, A., and Dano, K., (1998), *Biochemistry*, **37(11)**:3612-3622

Ploug, M., Rahbek-Nielsen, H., Ellis, V., Roepstorff, P., and Dano, K., (1995), *Biochemistry*, **34(39)**:12524-12534

Steiner, D., (1998), *Current Opinion in Chemical Biology*, **2**:31-39

Stephens, R.W., Bokman, A.M., Myohanen, H.T., Reisberg, T., Tapiovaara, H., Pedersen, N., Grondahl-Hansen, J., Llinas, M., and Vaheri, A., (1992), *Biochemistry*, **31**:7572-7579

Terrett, N., (2000), *Drug Discov Today*, **5(5)**:211-212

- 95 -

CLAIMS:

1. A method of facilitating production of a protein or derivative thereof from a negative sense single stranded RNA virus, said method comprising expressing in a host cell a nucleic acid molecule encoding said protein or derivative thereof, the nucleotide sequence of which nucleic acid molecule is optimised for expression by a eukaryotic cell.
2. The method according to claim 1 wherein said virus is a virus from the family Paramyxoviridae.
3. The method according to claim 2 wherein said virus is of the sub-family Pneumovirinae.
4. The method according to claim 3 wherein said virus is respiratory syncytial virus.
5. The method according to any one of claims 1-4 wherein said protein directly or indirectly facilitates fusion of any one or more viral components with any one or more host cells components.
6. The method according to claim 5 wherein said protein is a F protein or derivative thereof.
7. The method according to claim 6 wherein said derivative is the F_{so1} fragment.
8. The method according to claim 5 wherein said protein is an N protein or derivative thereof.
9. The method according to claim 5 wherein said protein is a P protein or derivative thereof.

- 96 -

10. The method according to claim 5 wherein said protein is a SH protein or derivative thereof.
11. The method according to any one of claims 1-10 wherein said eukaryotic host cell is a mammalian cell.
12. The method according to claim 11 wherein said mammalian cell is a 293 cell.
13. The method according to claim 11 wherein said mammalian cell is a Chinese Hamster Ovary Cell.
14. The method according to any one of claims 11-13 wherein said optimisation is codon optimisation and/or nucleotide splice site deletion.
15. The method according to claim 14, wherein said codon optimisation comprises modification of at least one A and/or T comprising codon to express G and C, respectively and said splice site deletion comprises deletion of at least one RNA splice site.
16. The method according to claim 14 or 15 wherein said optimised protein encoding nucleic acid molecule further comprises one or more endonuclease restriction sites.
17. The method according to any one of claims 14-16 wherein said optimised F protein encoding nucleic acid sequence corresponds to the sequence defined by <400>3 or derivative thereof.
18. The method according to any one of claims 14-16 wherein said optimised F protein encoding nucleic acid sequence corresponds to the sequence defined by <400>5 or derivative thereof.

- 97 -

19. The method according to any one of claims 14-16 wherein said optimised F_{SOL} protein encoding nucleic acid sequence corresponds to the sequence defined by <400>4 or derivative thereof.

20. The method according to any one of claims 14-16 wherein said optimised F_{SOL} protein encoding nucleic acid sequence corresponds to the sequence defined by <400>6 or derivative thereof.

21. The method according to any one of claims 14-16 wherein said optimised P protein encoding nucleic acid sequence corresponds to the sequence defined by <400>556 or derivative thereof.

22. The method according to any one of claims 14-16 wherein said optimised N protein encoding nucleic acid sequence correspond to the sequence defined by <400>559 or derivative thereof.

23. The method according to any one of claims 14-16 wherein said SH protein encoding nucleic acid sequence corresponds to the sequence defined by <400>562 or derivative thereof.

24. An optimised nucleic acid molecule or derivative thereof as described in any one of claims 1-23.

25. A protein molecule encoded by the optimised nucleic acid molecule of claim 24 or derivative, equivalent, analogue or mimetic thereof.

26. A method of regulating the functional activity of a viral F protein, which protein in its non-fully functional form comprises an F2 portion linked, bound or otherwise associated with an F1 portion, which F2 portion comprises an intervening peptide sequence, said method comprising modulating cleavage of said intervening peptide sequence wherein excision of at least part of said intervening sequence from said non-fully

- 98 -

functional form of said protein up-regulates F protein functional activity.

27. The method according to claim 26 wherein said method comprises expressing in a host cell a nucleic acid molecule encoding said protein or derivative thereof, the nucleotide sequence of which nucleic acid molecule is optimised for expression by a eukaryotic cell.

28. The method according to claim 27 wherein said virus is a virus from the family Paramyxoviridae.

29. The method according to claim 28 wherein said virus is of the sub-family Pneumovirinae.

30. The method according to any one of claims 26-29 wherein said cleavage events occur at the cleavage sites defined by the peptide sequence RARR (<400>564) and KKRKR (<400>563).

31. The method according to any one of claims 26-29 wherein said F protein, in its non-fully functional form, comprises the structure:

$$X_1 X_2 X_3$$

wherein:

X_1 comprises the non-intervening peptide sequence region of the F_2 portion;

X_2 comprises the intervening peptide sequence region of the F_2 portion; and

X_3 comprises the F_1 portion

32. The method according to claim 31 wherein said cleavage events occur at the cleavage sites defined by the peptide sequence RARR (<400>564) and KKRKR (<400>563).

- 99 -

33. The method according to any one of claims 26-32 wherein said regulation is down-regulation.

34. A method for detecting an agent capable of regulating the functional activity of a viral F protein or derivative thereof said method comprising contacting a eukaryotic cell expressing an optimised nucleic acid molecule in accordance with the method of any one of claims 1-23 with a putative modulatory agent and detecting an altered expression phenotype and/or functional activity.

35. The method of claim 34 wherein said viral F protein is a non-fully functional form of said protein and wherein said agent modulates cleavage of the intervening peptide sequence.

36. A method for detecting an agent capable of regulating the functional activity of a viral F protein or derivative thereof said method comprising contacting a host cell, which host cell expresses a nucleic acid molecule encoding the non-fully functional form of said viral F protein derivative thereof, with a putative modulatory agent and detecting an altered expression phenotype and/or altered functional activity wherein said agent modulates the site 2 cleavage event.

37. A method for analysing, designing and/or modifying an agent capable of interacting with a viral F protein or derivative thereof and modulating at least one functional activity associated with said protein, which protein is produced in accordance with the method of any one of claims 1-23, said method comprising contacting said F protein or derivative thereof with a putative agent and assessing the degree of interactive complementarity of said agent with said protein.

38. The method of claim 37 wherein said virus is a virus from the family Paramyxoviridae.

- 100 -

39. The method according to claim 38 wherein said virus is of the sub-family Pneumovirinae.
40. The method according to claim 39 wherein said virus is respiratory syncytial virus.
41. An agent capable of interacting with a viral F protein and modulating at least one functional activity associated with said viral protein.
42. The agent according to claim 41 wherein said agent is an antagonist which interacts with a sequence selected from:

CFASGQNITE <400>9	FASGQNITEE <400>10
ASGQNITEEF <400>11	SGQNITEEFY <400>12
GQNITEEFYQ <400>13	QNITEEFYQS <400>14
NITEEFYQST <400>15	ITEEFYQSTC <400>16
TEEFYQSTCS <400>17	EEFYQSTCSA <400>18
EFYQSTCSAV <400>19	FYQSTCSAVS <400>20
YQSTCSAVSK <400>21	QSTCSAVSKG <400>22
STCSAVSKGY <400>23	TCSAVSKGYL <400>24
CSAVSKGYLS <400>25	SAVSKGYLSA <400>26
AVSKGYLSAL <400>27	VSKGYLSALR <400>28
SKGYLSALRT <400>29	KGYSALRTG <400>30
GYLSALRTGW <400>31	YLSALRTGWY <400>32
LSALRTGWYT <400>33	SALRTGWYTS <400>34
ALRTGWYTSV <400>35	LRTGWYTSVI <400>36
RTGWYTSVIT <400>37	TGWYTSVITI <400>38
GWYTSVITIE <400>39	WYTSVITIEL <400>40
YTSVITIELS <400>41	TSVITIELSN <400>42
SVITIELSNI <400>43	VITIELSNIK <400>44
ITIELSNIKK <400>45	TIELSNIKKN <400>46
IELSNIKKNK <400>47	ELSNIKKNKC <400>48
LSNIKKNKC <400>49	SNIKKNKCNG <400>50
NIKKNKCNGT <400>51	IKKNKCNGTD <400>52
KKNKCNGTDA <400>53	KNKCNGTDAK <400>54
NKCNGTDAKV <400>55	KCNGTDAKVK <400>56
CNGTDAKVKL <400>57	NGTDAKVCLI <400>58
GTDAAVKLIK <400>59	TDAAVKLIKQ <400>60
DAVKLIKQEL <400>61	AKVKLIKQEL <400>62
KVKLIKQELD <400>63	VKLIKQELDK <400>64
KLIKQELDKY <400>65	LKQELDKYK <400>66
IKQELDKYKN <400>67	KQELDKYKNA <400>68
QELDKYKNAV <400>69	ELDKYKNAV <400>70
LDKYKNAVTE <400>71	DKYKNAVTEL <400>72
KYKNAVTELO <400>73	YKNAVTELOL <400>74
KNAVTELOLL <400>75	NAVTELOLLM <400>76
AVTELOLLMQ <400>77	VTELOLLMQS <400>78

- 101 -

TELQLLMQST	<400>79	ELQLLMQSTQ	<400>80
LQLLMQSTQA	<400>81	QLLMQSTQAT	<400>82
LLMQSTQATN	<400>83	LMQSTQATNN	<400>84
MQSTQATNNR	<400>85	QSTQATNNRA	<400>86
STQATNNRAR	<400>87	TQATNNRARR	<400>88
QATNNRARRE	<400>89	ATNNRARREL	<400>90
TNNRARRELP	<400>91	NNRARRELPR	<400>92
NNRARRELPRF	<400>93	RARRELPRFM	<400>94
ARRELPRFMN	<400>95	RRELPRFMNY	<400>96
RELPRFMNYT	<400>97	ELPRFMNYTL	<400>98
LPRFMNYTLN	<400>99	PRFMNYTLNN	<400>100
RFMNYTLNNA	<400>101	FMNYTLNNAK	<400>102
MNYTLNNAKK	<400>103	NYTLNNAKKT	<400>104
YTLNNAKKTN	<400>105	TLNNAKKTNV	<400>106
LNNAKKTNTV	<400>107	NNAKKTNTVL	<400>108
NAKKTNTVLS	<400>109	AKKTNTVLSK	<400>110
KKTNTVLSKK	<400>111	KTNVLSKKR	<400>112
TNVLSKKRK	<400>113	NVLSKKRKR	<400>114
VTLSKKRRR	<400>115	TLSKKRKRRF	<400>116
LSKKRKRRL	<400>117	SKKRKRRLG	<400>118
KKRKRRLGF	<400>119	KRKRRFLGFL	<400>120
RKRRLGFL	<400>121	KRRFLGFLG	<400>122
RRFLGFLGV	<400>123	RFLGFLGVG	<400>124
FLGFLGVGS	<400>125	LGFLGVGSA	<400>126
GFLGVGSAI	<400>127	FLLGVGSAIA	<400>128
LLGVGSAIAS	<400>129	LGVGSAIASG	<400>130
GVGSAIASGV	<400>131	VGSIAASGVA	<400>132
GSAIASGVAV	<400>133	SAIASGVAVS	<400>134
AIASGVAVSK	<400>135	IASGVAVSKV	<400>136
ASGVAVSKVL	<400>137	SGVAVSKVLH	<400>138
GVAVSKVLHL	<400>139	VAVSKVLHLE	<400>140
AVSKVLHLEG	<400>141	VSKVLHLEGE	<400>142
SKVLHLEGEV	<400>143	KVLHLEGEVN	<400>144
VLHLEGEVNK	<400>145	LHLEGEVNKI	<400>146
HLEGEVNKIK	<400>147	LEGEVNKIKS	<400>148
EGEVNKIKA	<400>149	GEVNKIKSAL	<400>150
EVNKIKSALL	<400>151	VNKIKSALLS	<400>152
NKIKSALLST	<400>153	KIKSALLSTN	<400>154
IKSALLSTNK	<400>155	KSALLSTNKA	<400>156
SALLSTNKAV	<400>157	ALLSTNKAVV	<400>158
LLSTNKAVVS	<400>159	LSTNKAVVSL	<400>160
STNKAVVSLS	<400>161	TNKAVVSLSN	<400>162
NKAVVSLSNG	<400>163	KAVVSLSNGV	<400>164
AVVSLSNGVS	<400>165	VVSLSNGVSV	<400>166
VSLSNGVSVL	<400>167	SLSNGVSVLT	<400>168
LSNGVSVLTS	<400>169	SNGVSVLTSK	<400>170
NGVSVLTSKV	<400>171	GVSVLTSKVL	<400>172
VSVLTSKVLD	<400>173	SVLTSKVLDL	<400>174
VLTSKVLDLK	<400>175	LTSKVLDLKN	<400>176
TSKVLDLKNY	<400>177	SKVLDLKNYI	<400>178
KVLDLKNYID	<400>179	VLDLKNYIDK	<400>180
LDLKNYIDKQ	<400>181	DLKNYIDKQL	<400>182
LKNYIDKQLL	<400>183	KNYIDKQLLP	<400>184
NYIDKQLLPI	<400>185	YIDKQLLPIV	<400>186

- 102 -

IDKQLLPIVN <400>187
KQLLPIVNKQ <400>189
LLPIVNBKQSC <400>191
PIVNBKQSCSI <400>193
VNBKQSCSISN <400>195
KQSCSISNIE <400>197
SCSISNIETV <400>199
SISNIETVIE <400>201
SNIETVIEFQ <400>203
IETVIEFQQK <400>205
TVIEFQQKNN <400>207
IEFQQKNNRL <400>209
FQQKNNRLLE <400>211
QKNNRLLEIT <400>213
NNRLLEITRE <400>215
RLLEITREFS <400>217
LEITREFSVN <400>219
ITREFSVNAG <400>221
REFSVNAGVT <400>223
FSVNAGVTP <400>225
VNAGVTPVST <400>227
AGVTPVSTY <400>229
VTPVSTYML <400>231
TPVSTYMLTN <400>233
VSTYMLTNSE <400>235
TYMLTNSELL <400>237
MLTNSELLSL <400>239
TNSELLSLIN <400>241
SELLSLINDM <400>243
LLSLINDMPI <400>245
SLINDMPITN <400>247
INDMPITNDQ <400>249
DMPITNDQKK <400>251
PITNDQKKLM <400>253
TNDQKKLMSN <400>255
DQKKLMSNNV <400>257
KKLMSNNVQI <400>259
LMSNNVQIVR <400>261
SNNVQIVRQQ <400>263
NVQIVRQQSY <400>265
QIVRQQSYSI <400>267
VRQQSYSIMS <400>269
QQSYSIMSII <400>271
SYSIMSIIKE <400>273
SIMSIIKEEV <400>275
MSIIKEEVLA <400>277
IIKEEVLAYV <400>279
KEEVLAYVVQ <400>281
EVLAYVVQLP <400>283
LAYVVQLPLY <400>285
YVVQLPLYGV <400>287
VQLPLYGVID <400>289
LPLYGVIDTP <400>291
LYGVIDTPCW <400>293

DKQLLPIVNK <400>188
QLLPIVNBKQS <400>190
LPIVNBKQSCS <400>192
IVNBKQSCSIS <400>194
NBKQSCSISNI <400>196
QSCSISNIET <400>198
CSISNIETVI <400>200
ISNIETVIEF <400>202
NIETVIEFQQ <400>204
ETVIEFQQKN <400>206
VIEFQQKNNR <400>208
EFQQKNNRLL <400>210
QQKNNRLLLEI <400>212
KNNRLLLEITR <400>214
NRLLEITREF <400>216
LLEITREFSV <400>218
EITREFSVNA <400>220
TREFSVNAGV <400>222
EFSVNAGVTT <400>224
SVNAGVTPV <400>226
NAGVTPVST <400>228
GVTPVSTYM <400>230
TTPVSTYMLT <400>232
PVSTYMLTNS <400>234
STYMLTNSEL <400>236
YMLTNSELLS <400>238
LTNSELLSLI <400>240
NSELLSLIND <400>242
ELLSLINDMP <400>244
LSLINDMPIT <400>246
LINDMPITND <400>248
NDMPITNDQK <400>250
MPITNDQKKL <400>252
ITNDQKKLMS <400>254
NDQKKLMSNN <400>256
QKKLMSNNVQ <400>258
KLMSNNVQIV <400>260
MSNNVQIVRQ <400>262
NNVQIVRQQS <400>264
VQIVRQQSYS <400>266
IVRQQSYSIM <400>268
RQQSYSIMSI <400>270
QSYSIMSIIK <400>272
YSIMSIIKEE <400>274
IMSIIKEEVL <400>276
SIIKEEVLAY <400>278
IKEEVLAYVV <400>280
EEVLAYVVQL <400>282
VLAYVVQLPL <400>284
AYVVQLPLYG <400>286
VVQLPLYGVI <400>288
QLPLYGVIDT <400>290
PLYGVIDTPC <400>292
YGVIDTPCWK <400>294

- 103 -

GVIDTPCWKL <400>295
IDTPCWKLHT <400>297
TPCWKLHTSP <400>299
CWKLHTSPLC <400>301
KLHTSPLCTT <400>303
HTSPLCTTNT <400>305
SPLCTTNTKE <400>307
LCTTNTKEGS <400>309
TTNTKEGSNI <400>311
NTKEGSNICL <400>313
KEGSNICLTR <400>315
GSNICLTRTD <400>317
NICLTRTRDRG <400>319
CLTRTRDRGWY <400>321
TRTRDRGWYCD <400>323
TDRGWYCDNA <400>325
RGWYCDNAGS <400>327
WYCDNAGSVS <400>329
CDNAGSVSFF <400>331
NAGSVSFFPQ <400>333
GSVSFFPQAE <400>335
VSFFPQAECK <400>337
FFPQAECKVQ <400>339
PQAECKVQSN <400>341
AETCKVQSNR <400>343
TCKVQSNRVF <400>345
KVQSNRVFCD <400>347
QSNRVFCDTM <400>349
NRVFCDTMNS <400>351
VFCDTMNSLT <400>353
CDTMNSLTLP <400>355
TMNSLTLPSE <400>357
NSLTLPSEVN <400>359
LTLPSEVNLC <400>361
LPSEVNLCNV <400>363
SEVNLCNVDI <400>365
VNLCNVDIFN <400>367
LCNVDIFNPK <400>369
NVDIFNPKYD <400>371
DIFNPKYDCK <400>373
FNPKYDCKIM <400>375
PKYDCKIMTS <400>377
YDCKIMTSKT <400>379
CKIMTSKTDV <400>381
IMTSKTDVSS <400>383
TSKTDVSSSV <400>385
KTDVSSSVIT <400>387
DVSSSVITSL <400>389
SSSVITSLGA <400>391
SVITSLGAIV <400>393
ITSLGAIVSC <400>395
SLGAIVSCYG <400>397
GAIVSCYGKT <400>399
IVSCYGKTKC <400>401

VIDTPCWKLH <400>296
DTPCWKLHTS <400>298
PCWLHTSPL <400>300
WKLHTSPLCT <400>302
LHTSPLCTTN <400>304
TSPLCTTNTK <400>306
PLCTTNTKEG <400>308
CTTNTKEGSN <400>310
TNTKEGSNIC <400>312
TKEGSNICLT <400>314
EGSNICLTRT <400>316
SNICLTRTRD <400>318
ICLTRTRDRGW <400>320
LTRTRDRGWYC <400>322
RTDRGWYCDN <400>324
DRGWYCDNAG <400>326
GWYCDNAGSV <400>328
YCDNAGSVSF <400>330
DNAGSVSFFP <400>332
AGSVSFFPQA <400>334
SVSFFPQAE <400>336
SFFPQAECK <400>338
FPQAECKVQ <400>340
QAECKVQSN <400>342
ETCKVQSNRV <400>344
CKVQSNRVFC <400>346
VQSNRVFCDT <400>348
SNRVFCDTMN <400>350
RVFCDTMNSL <400>352
FCDTMNSLTL <400>354
DTMNSLTLP <400>356
MNSLTLPSEV <400>358
SLTLPSEVNL <400>360
TLPSEVNLCN <400>362
PSEVNLCNVD <400>364
EVNLCNVDIF <400>366
NLCNVDIFNP <400>368
CNVDIFNPKY <400>370
VDIFNPKYDC <400>372
IFNPKYDCKI <400>374
NPKYDCKIMT <400>376
KYDCKIMTSK <400>378
DCKIMTSKTD <400>380
KIMTSKTDVS <400>382
MTSKTDVSSS <400>384
SKTDVSSSVI <400>386
TDVSSSVITS <400>388
VSSSVITSLG <400>390
SSVITSLGAI <400>392
VITSLGAIVS <400>394
TSLGAIVSCY <400>396
LGAIVSCYCK <400>398
AIVSCYGKTK <400>400
VSCYGKTKCT <400>402

- 104 -

SCYGKTKCTA	<400>403	CYGKTKCTAS	<400>404
YGKTKCTASN	<400>405	GKTKCTASNK	<400>406
KTKCTASNKN	<400>407	TKCTASNKNR	<400>408
KCTASNKNRG	<400>409	CTASNKNRGI	<400>410
TASNKNRGII	<400>411	ASNKNRGIK	<400>412
SNKNRGIKT	<400>413	NKNRGIKTF	<400>414
KNRGIKTF	<400>415	NRGIKTF	<400>416
RGIIKTF	<400>417	GIKTF	<400>418
IIKTF	<400>419	IKTF	<400>420
KTFSNGCDY	<400>421	TFSNGCDYVS	<400>422
FSNGCDYVSN	<400>423	SNGCDYVSNK	<400>424
NGCDYVSNKG	<400>425	GCDYVSNKGV	<400>426
CDYVSNKGV	<400>427	DYVSNKGVDT	<400>428
YVSNKGVDTV	<400>429	VSNKGVDTV	<400>430
SNKGVDTVSV	<400>431	NKGVDTVSVG	<400>432
KGVDTVSVGN	<400>433	GVDTVSVGNT	<400>434
VDTVSVGNTL	<400>435	DTVSVGNTLY	<400>436
TVSVGNTLYY	<400>437	VSVGNTLYYV	<400>438
SVGNTLYYVN	<400>439	VGNTLYYVNK	<400>440
GNTLYYVNBK	<400>441	NTLYYVNBKQ	<400>442
TLYYVNBKQ	<400>443	LYYVNBKQEG	<400>444
YYVNBKQEG	<400>445	YVNBKQEGK	<400>446
VNBKQEGKSL	<400>447	NKQEGKSLYV	<400>448
KQEGKSLYV	<400>449	QEGKSLYVKG	<400>450
EGKSLYVKG	<400>451	GKSLYVKGEP	<400>452
KSLYVKGEP	<400>453	SLYVKGEP	<400>454
LYVKGEP	<400>455	YVKGEP	<400>456
VKGEP	<400>457	KGEP	<400>458
GEP	<400>459	EPI	<400>460
PI	<400>461	I	<400>462
INFYDPLV	<400>463	NFYDPLVFP	<400>464
FYDPLVFP	<400>465	YDPLVFP	<400>466
DPLVFP	<400>467	PLVFP	<400>468
LVFP	<400>469	VFP	<400>470
FPSDEFDASI	<400>471	PSDEFDASIS	<400>472
SDEFDASISQ	<400>473	DEFDASISQV	<400>474
EFDASISQV	<400>475	FDASISQVNE	<400>476
DASISQVNEK	<400>477	ASISQVNEKI	<400>478
SISQVNEKIN	<400>479	ISQVNEKINQ	<400>480
SQVNEKINQS	<400>481	QVNEKINQSL	<400>482
VNEKINQSLA	<400>483	NEKINQSLAF	<400>484
EKINQSLAFI	<400>485	KINQSLAFIR	<400>486
INQSLAFIRK	<400>487	NQSLAFIRKS	<400>488
QSLAFIRKSD	<400>489	SLAFIRKSDE	<400>490
LAFIRKSDEL	<400>491	AFIRKSDELL	<400>492
FIRKSDELLH	<400>493	IRKSDELLHN	<400>494
RKSDELLHNV	<400>495	KSDELLHNVN	<400>496
SDELLHNVNA	<400>497	DELLHNVNAG	<400>498
ELLHNVNAGK	<400>499	LLHNVNAGKS	<400>500
LHNVNAGKST	<400>501	HNVNAGKSTT	<400>502
NVNAGKSTTN	<400>503	VNAGKSTTNI	<400>504
NAGKSTTNIM	<400>505	AGKSTTNIMI	<400>506
GKSTTNIMIT	<400>507	KSTTNIMITT	<400>508
STTNIMITTI	<400>509	TTNIMITTII	<400>510

- 105 -

TNIMITTIII <400>511	NIMITTTIIIV <400>512
IMITTIIIVI <400>513	MITTTIIIVII <400>514
ITTTIIIVIIV <400>515	TTTIIIVIIVI <400>516
TTTIIIVIIVIL <400>517	IIIIVIIVILL <400>518
IIVIIVILLS <400>519	IVIIVILLSL <400>520
VIIIVILLSLI <400>521	IIVILLSLIA <400>522
IVILLSLIAV <400>523	VILLSLIAVG <400>524
ILLSLIAVGL <400>525	LLSLIAVGLL <400>526
LSLIAVGLLL <400>527	SLIAVGLLLY <400>528
LIAVGLLLYC <400>529	IAVGLLLYCK <400>530
AVGLLLYCKA <400>531	VGLLLYCKAR <400>532
GLLYCKARS <400>533	LLLYCKARST <400>534
LLYCKARSTP <400>535	LYCKARSTPV <400>536
YCKARSTPVT <400>537	CKARSTPVTL <400>538
KARSTPVTLS <400>539	ARSTPVTLSK <400>540
RSTPVTLSKD <400>541	STPVTLSKDQ <400>542
TPVTLSKDQL <400>543	PVTLSKDQLS <400>544
VTLSDQLSG <400>545	TLSDQLSGI <400>546
LSKDQLSGIN <400>547	SKDQLSGINN <400>548
KDQLSGINNI <400>549	DQLSGINNIA <400>550
QLSGINNIAF <400>551	LSGINNIAFS <400>552
SGINNIAFSN <400>553	

43. The agent according to claim 42 wherein said antagonist interacts with a sequence selected from <400>88, <400>89, <400>90, <400>91, <400>92, <400>93 or <400>94.

44. A viral F protein variant comprising a mutation in the intervening peptide sequence wherein said variant exhibits modulated functional activity relative to wild-type F protein or a derivative, homologue, analogue, chemical equivalent or mimetic of said variant.

45. The variant according to claim 44 wherein said variant exhibits down-regulated functional activity relative to wild-type F protein.

46. The variant according to claim 44 or claim 45 wherein said virus is a virus from the family Paramyxoviridae.

47. The variant according to claim 46 wherein said virus is of the sub-family Pneumovirinae.

48. The variant according to claim 47 wherein said virus is respiratory syncytial virus.

- 106 -

49. The variant according to claim 48 wherein said variant comprises a mutation in the cleavage site defined by amino acids RARR (<400>564).

50. The variant according to claim 49 wherein said mutation comprises one or more of the amino acid substitutions selected from the following list:

- (i) R106G
- (ii) A107Q
- (iii) R108G.

51. The variant according to claim 50 wherein said variant comprises the sequence substantially as set forth in <400>565.

52. The variant according to any one of claims 44-48 wherein said variant comprises a multiple amino acid deletion from the intervening peptide sequence.

53. The variant according to claim 52 wherein said amino acid deletion is a partial deletion of the intervening peptide sequence.

54. The variant according to claim 53 wherein said deletion is a deletion of the peptide sequence

RARRELPRFMNYTLNNAKKTNVTLS <400>569

55. The variant according to claim 54 wherein said variant comprises the amino acid sequence substantially as set forth in <400>567.

56. An isolated nucleic acid molecule selected from the list consisting of:

- 107 -

- (i) An isolated nucleic acid molecule or derivative or equivalent thereof comprising a nucleotide sequence encoding or complementary to a sequence encoding a viral F protein variant or derivative, homologue, analogue, chemical equivalent or mimetic of said variant, which variant comprises a mutation in the intervening peptide sequence wherein said variant exhibits modulated functional activity relative to wild-type F protein.
- (ii) An isolated nucleic acid molecule or derivative or equivalent thereof comprising a nucleotide sequence encoding or complementary to a sequence encoding a viral F protein variant or derivative, homologue, analogue, chemical equivalent or mimetic of said variant, which variant comprises a mutation in the intervening peptide sequence wherein said variant exhibits down-regulated functional activity relative to wild-type F protein.
- (iii) An isolated nucleic acid molecule or derivative or equivalent thereof comprising a nucleotide sequence encoding or complementary to a sequence encoding a respiratory syncytial virus F protein or derivative, homologue, analogue, chemical equivalent or mimetic of said variant, which variant comprises a mutation in the cleavage site defined by amino acids RARR wherein said variant exhibits down-regulated functional activity relative to wild-type F protein.
- (iv) An isolated nucleic acid molecule or derivative or equivalent thereof comprising a nucleotide sequence encoding or complementary to a sequence encoding a respiratory syncytial virus F protein variant or derivative, homologue, analogue, chemical equivalent or mimetic of said variant, which variant comprises one or more of the amino acid substitutions selected from the following list:

- 108 -

- (a) R106G
 - (b) A107Q
 - (c) R108G
- (v) An isolated nucleic acid molecule or derivative or analogue thereof comprising a nucleotide sequence encoding or complementary to a sequence encoding a viral F protein variant or derivative, homologue, analogue, chemical equivalent or mimetic of said variant, which variant comprises a multiple amino acid deletion from the intervening peptide sequence wherein said variant exhibits down-regulated functional activity relative to wild-type F protein.
- (vi) An isolated nucleic acid molecule or derivative or analogue thereof comprising a nucleotide sequence encoding or complementary to a sequence encoding a viral F protein variant or derivative, homologue, analogue, chemical equivalent or mimetic of said variant, which variant comprises a partial deletion of the intervening peptide sequence and more preferably a deletion of the peptide sequence
- RARRELPRFMNYTLNNAKKTNVTLS <400>569.
- (vii) An isolated nucleic acid molecule or derivative or analogue thereof comprising a nucleotide sequence encoding or complementary to a sequence encoding a viral F protein variant or derivative, homologue, analogue, chemical equivalent or mimetic of said variant, which variant comprises the amino acid sequence substantially as set forth in <400>567.
- (viii) An isolated nucleic acid molecule or derivative or analogue thereof comprising a nucleotide sequence encoding or complementary to a sequence encoding a viral F protein variant or derivative, homologue,

- 109 -

analogue, chemical equivalent or mimetic of said variant, which variant comprises the amino acid sequence substantially as set forth in <400>565.

(ix) An isolated nucleic acid molecule or derivative or analogue thereof comprising the nucleotide substantially as set forth in <400>568.

(x) An isolated nucleic acid molecule or derivative or analogue thereof comprising the nucleotide substantially as set forth in <400>566.

57. The isolated nucleic acid molecule of claim 56 wherein said virus is a virus from the family Paramyxoviridae.

58. The isolated nucleic acid molecule of claim 57 wherein said virus is of the sub-family Pneumovirinae.

59. The isolated nucleic acid molecule of claim 58 wherein said virus is respiratory syncytial virus.

60. A recombinant viral construct comprising a nucleic acid molecule encoding a viral F protein or derivative thereof, the nucleotide sequence of which nucleic acid molecule comprises codons optimised for expression in a eukaryotic cell, wherein said recombinant viral construct is effective in inducing, enhancing or otherwise stimulating an immune response to said F protein.

61. A recombinant viral construct comprising a nucleic acid molecule encoding a viral F protein variant or derivative thereof wherein said recombinant viral construct is effective in inducing, enhancing or otherwise stimulating an immune response to said F protein variant.

62. A vaccine comprising a recombinant viral construct which construct comprises a nucleic acid molecule encoding a respiratory syncytial virus F protein or derivative thereof,

- 110 -

the nucleic sequence of which nucleic acid molecule is optimised for expression in a eukaryotic cell wherein said recombinant viral construct is effective in inducing, enhancing or otherwise stimulating an immune response to said F protein.

63. A vaccine comprising a recombinant viral construct which construct comprises a nucleic acid molecule encoding a respiratory syncytial virus F protein variant or derivative thereof, wherein said recombinant viral construct is effective in inducing, enhancing or otherwise stimulating an immune response to said F protein variant.

64. A vaccine according to claim 62 or claim 63 wherein said nucleotide sequence is defined in one of <400>5, <400>6, <400>566 or <400>568.

65. Use of the agent according to any one of claims 41-43 or identified in accordance with the method of any one of claims 34-40 to modulate F protein functional activity.

66. Use of the agent according to any one of claims 41-43 or identified in accordance with the method of any one of claims 34-40 in the therapeutic and/or prophylactic treatment of conditions characterised by infection with a negative sense single stranded RNA virus.

67. A method of modulating at least one functional activity associated with a viral F protein in a subject, said method comprising introducing into said subject an effective amount of a F protein modulatory agent according to any one of claims 41-43 or identified in accordance with the method of any one of claims 34-40 for a time and under conditions sufficient for said agent to interact with said F protein.

68. The method according to claim 68 wherein said functional activity is F protein mediated host cell virion fusion and/or virion budding and said modulating is down-regulation.

- 111 -

69. A method of modulating at least one functional activity associated with a viral F protein, said method comprising contacting said viral F protein with an effective amount of a F protein modulatory agent according to any one of claims 41-43 or identified in accordance with the method of any one of claims 34-40 for a time and under conditions sufficient for said agent to interact with said F protein.

70. A method for the treatment and/or prophylaxis of a condition characterised by infection with a negative sense single stranded RNA virus in a subject, said method comprising administering to said subject an effective amount of an agent according to any one of claims 41-43 or identified in accordance with the method of any one of claims 34-40 which agent is capable of down-regulating at least one functional activity of the F protein expressed by said virus, for a time and under conditions sufficient for said agent to interact with said F protein.

71. A method for the treatment and/or prophylaxis of a condition characterised by infection with a negative sense single stranded virus in a subject, said method comprising administering to said subject an effective amount of a composition comprising an F protein or derivative thereof, F protein variant or derivative thereof and/or a nucleic acid molecule encoding said F protein or F protein variant or a derivative, homologue, analogue, chemical equivalent or mimetic of said protein or nucleic acid molecule for a time and under conditions sufficient for said composition to down regulate said viral F protein functional activity.

72. The method according to claim 71 wherein said subject is a mammal.

73. The method according to claim 72 wherein said mammal is a human.

74. Use of an agent capable of modulating at least one functional activity of a viral F protein which agent according to any one of claims 41-43 or identified in accordance with the method of any one of claims 34-40 in the manufacture of a medicament for the treatment and/or prophylaxis of a condition characterised by infection with negative sense

- 112 -

single stranded RNA virus.

75. Use of a composition comprising an F protein or derivative thereof, F protein variant or derivative thereof, nucleic acid molecule encoding said F protein or F protein variant according to any one of claims 41-43 or identified in accordance with the method of any one of claims 34-40 or a derivative, homologue, analogue, chemical equivalent or mimetic of said protein or nucleic acid molecule, in the manufacture of a medicament for the treatment and/or prophylaxis of a condition characterised by infection with a negative sense single stranded RNA virus.

76. Use of an agent, which agent, according to any one of claims 41-43 or identified in accordance with the method of any one of claims 34-40 in the manufacture of a medicament for the modulation of at least one viral F protein associated functional activity.

77. Agents for use in modulating the functional activity of a viral F protein wherein said agent is identified in accordance with the method of any one of claims 34-40.

78. Agents for use in the treatment and/or prophylaxis of a condition characterised by infection with a negative sense single stranded RNA virus wherein said agent is identified in accordance with the methods of any one of claims 34-40.

79. A composition comprising an F protein or derivative thereof, F protein variant or derivative thereof, a nucleic acid molecule encoding said F protein or F protein variant or a derivative, homologue, analogue, chemical equivalent or mimetic of said protein or nucleic acid molecule for use in the treatment and/or prophylaxis of a condition characterised by infection with a negative sense single stranded RNA virus.

1/40

MELLILKANAIITILTAVTECFASGQNIITEEFYQSTCSAVSKGYLSALRTGWYTSVITIELSNIKKKNCN
GTDKVKLIKQELDKYKNAVTELOLLMQSTQATNNRARRELPRFMNYTLNNAKKTNTVTLSSKKRRRFLGF
LLGVGSAIASGVAVSKVLHLEGEVNIKSALLSTNKAVVSLSNVSVLTSKVLDLKNYIDKQLLPVVKQ
SCSISNIETVIEFQQKNNRLLLEITREFSVNAGVTPVSTYMLTNSSELLSLINDMPIITNDQKKLMSNNVQI
VRQQSYSIMSIIKEEVLAIVVQPLVYGVIDTPCWKLHTSPLCTNTKEGSNICLTRDRGWYCDNAGSVS
FFPQAETCKVQSNRVFCDTMNSLTLPSEVNLNVDFENPKYDCKIMTSKTDVSSSVITSLGAIVSCYGT
KCTASNKNRGIKTFNSNGCDYVSNKGVDTVSVGNTLYYVVKQEGKSLYVKGEPIINFYDPLVFPSEFFDA
SISQVNEKINQSLAFIRKSDELLHNVNAGKSTTNIMITTIIIVIIIVILLSLIAGVLLLYCKARSTPVTLS
KDQLSGINNIAFSN*

70

140

210

280

350

420

490

560

574

Figure 1a

2/40

MELLILKANAI TTILTA VTFCE FASGQNI TEEFYQSTCSAVSKGYLSALRTGWYTSVITIELSNIKKNCN 70
GTDKVKLIKQELDKYKNAVTE LQLLMQSTQATNNRARRRELPRFMNYT LNNAKTNNVTLSKKRRRFLGF 140
LLGVGSAIAGVAVSKVLHLEGEV NKIKSALLSTNKAVVSLSGVSVLT SKVLDLKNYIDKQLLP IVNKQ 210
SCSISNIETVIEFQOKNNRLL EITREFSVNAGVTPVSTYMLTNS ELLSLINDMPITNDQKKLMSNNVQI 280
VRQQSYSIMSIKEEV LAYVQPLPYGVIDTPCWKLHTSP LCTTNTKEGSNICLTRDRGWYCDNAGSVS 350
FFPQAE TCKVQSNRVFCDTMNSLTLPSEVNL CNVDIFNPKYDCKIMTSKTDVSSSVITSLGAI VSCYGT 420
KCTASNKNRGI IKTFSNGCDYVSNKGVDTVSVGNTLYYV NKQEGKSLYVKGEPIINFYDPLVFP SDEFDA 490
SISQVNEKINQSLAFIRKSDELLHNVNAGKSTN 524

Figure 1b

3/40

F and F_{sol} forms of the RSV fusion Protein

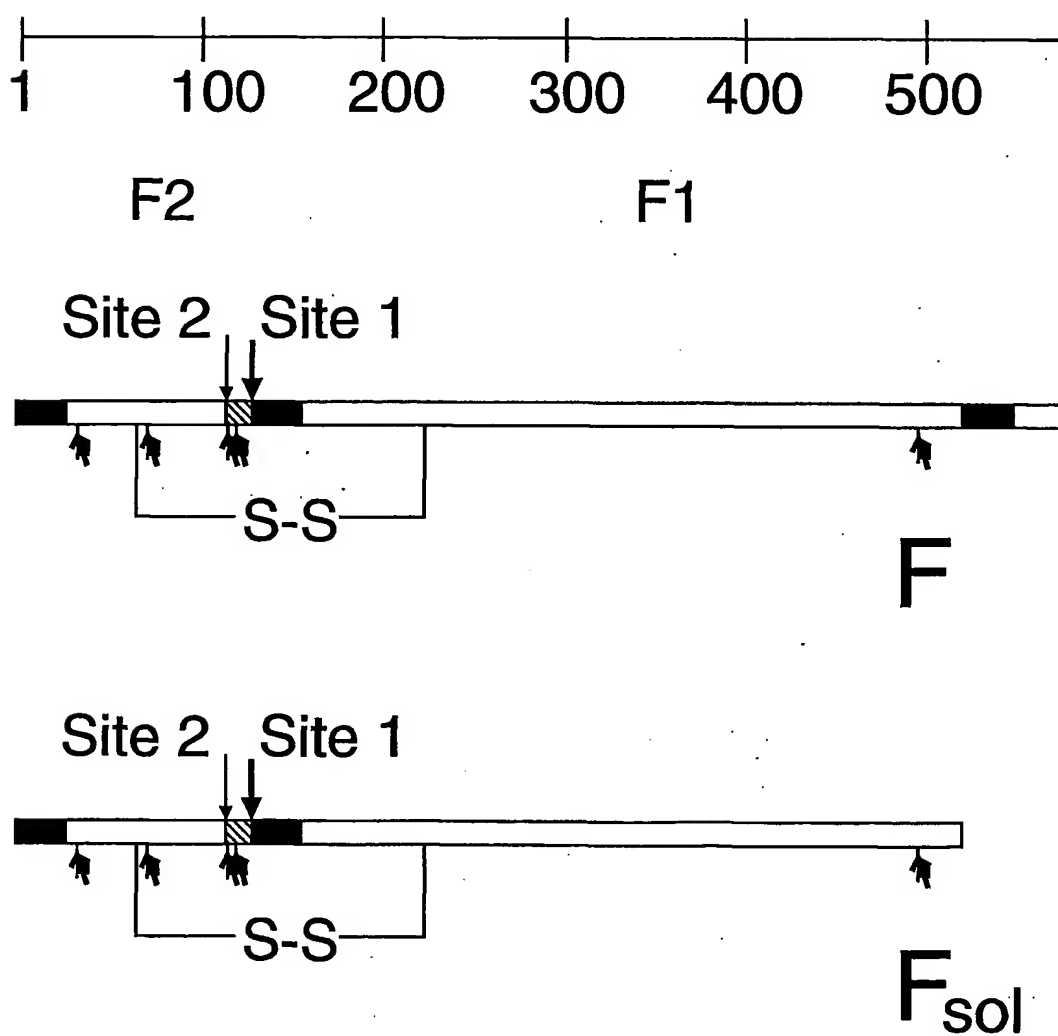


Figure 1c
SUBSTITUTE SHEET (RULE 26)

4/40

(i)
(ii)
(iii)
(iv)
(v)
(vi)
(vii)

Figure 2a
SUBSTITUTE SHEET (RULE 26)

5/40

ClustalW Formatted Alignments

		10	20	30																										
<i>F.viral</i>	A	T	G	G	A	G	T	T	G	C	T	A	A	T	C	C	T	C	A	A	A	G	C	A	A	A	T	G	C	A
<i>F</i>	A	T	G	G	A	G	T	T	G	C	T	A	A	T	C	C	T	C	A	A	A	G	C	A	A	A	T	G	C	A
<i>F.opt</i>	A	T	G	G	A	G	C	T	G	C	T	G	A	T	C	C	T	G	A	A	G	G	C	C	A	A	C	G	C	C
	A	T	G	G	A	G	T	T	G	C	T	A	A	T	C	C	T	C	A	A	A	G	C	A	A	A	T	G	C	A

		40	50	60																										
<i>F.viral</i>	A	T	T	A	C	C	A	C	A	A	T	C	C	T	C	A	C	T	G	C	A	G	T	C	A	C	A	T	T	T
<i>F</i>	A	T	T	A	C	C	A	C	A	A	T	C	C	T	C	A	C	T	G	C	G	G	T	C	A	C	C	T	T	T
<i>F.opt</i>	A	T	C	A	C	C	A	C	C	A	T	C	C	T	G	A	C	G	C	G	G	T	G	A	C	C	T	T	C	
	A	T	T	A	C	C	A	C	A	A	T	C	C	T	C	A	C	T	G	C	G	G	T	C	A	C	C	T	T	T

		70	80	90																										
<i>F.viral</i>	T	G	T	T	T	T	G	C	T	T	C	T	G	G	T	C	A	A	A	A	C	A	T	C	A	C	T	G	A	A
<i>F</i>	T	G	T	T	T	T	G	C	T	T	C	T	G	G	T	C	A	A	A	A	C	A	T	C	A	C	T	G	A	A
<i>F.opt</i>	T	G	C	T	T	C	G	C	C	T	C	T	G	G	C	C	A	G	A	A	C	A	T	C	A	C	T	G	A	G
	T	G	T	T	T	T	G	C	T	T	C	T	G	G	T	C	A	A	A	A	C	A	T	C	A	C	T	G	A	A

		100	110	120																										
<i>F.viral</i>	G	A	A	T	T	T	A	T	C	A	A	T	C	A	A	C	A	T	G	C	A	G	T	G	C	A	G	T	T	
<i>F</i>	G	A	A	T	T	T	A	T	C	A	A	T	C	A	A	C	A	T	G	C	A	G	T	G	C	A	G	T	T	
<i>F.opt</i>	G	A	G	T	T	C	T	A	C	C	A	G	A	G	C	A	C	T	T	G	T	T	C	C	G	C	T	G	T	G
	G	A	A	T	T	T	A	T	C	A	A	T	C	A	A	C	A	T	G	C	A	G	T	G	C	A	G	T	T	

		130	140	150																											
<i>F.viral</i>	A	G	C	A	A	A	G	G	C	T	A	T	C	T	T	A	G	T	G	C	T	C	T	G	A	G	A	A	C	T	
<i>F</i>	A	G	C	A	A	A	G	G	A	T	A	T	C	T	T	A	G	T	G	C	T	C	T	G	A	G	A	A	C	C	
<i>F.opt</i>	A	G	C	A	A	A	G	G	G	C	T	A	C	C	T	G	A	G	C	G	C	C	C	T	G	A	G	A	A	C	C
	A	G	C	A	A	A	G	G	C	T	A	T	C	T	T	A	G	T	G	C	T	C	T	G	A	G	A	A	C	C	

		160	170	180																										
<i>F.viral</i>	G	G	T	T	G	G	T	A	T	A	C	C	A	G	T	G	T	T	A	T	A	A	C	T	A	T	A	G	A	A
<i>F</i>	G	G	T	T	G	G	T	A	T	A	C	C	A	G	T	G	T	T	A	T	A	A	C	T	A	T	A	G	A	A
<i>F.opt</i>	G	G	T	T	G	G	T	A	C	A	C	C	A	G	C	G	T	G	A	T	C	A	C	C	A	T	C	G	A	G
	G	G	T	T	G	G	T	A	T	A	C	C	A	G	T	G	T	T	A	T	A	A	C	T	A	T	A	G	A	A

		190	200	210																										
<i>F.viral</i>	T	T	A	A	G	T	A	A	T	A	T	C	A	A	G	A	A	A	A	A	T	A	A	G	T	G	T	A	A	T
<i>F</i>	T	T	A	A	G	T	A	A	T	A	T	C	A	A	G	A	A	A	A	A	T	A	A	G	T	G	T	A	A	T
<i>F.opt</i>	C	T	G	A	G	C	A	A	C	A	T	C	A	A	G	A	A	G	A	A	C	A	A	G	T	G	C	A	A	C
	T	T	A	A	G	T	A	A	T	A	T	C	A	A	G	A	A	A	A	A	T	A	A	G	T	G	T	A	A	T

		220	230	240																										
<i>F.viral</i>	G	G	A	A	C	C	G	A	T	G	C	T	A	A	G	G	T	A	A	A	A	T	T	G	A	T	A	A	A	A
<i>F</i>	G	G	T	A	C	C	G	A	T	G	C	T	A	A	G	G	T	A	A	A	A	T	T	G	A	T	A	A	A	A
<i>F.opt</i>	G	G	C	A	C	C	G	A	C	G	C	C	A	A	G	G	T	G	A	A	G	C	T	G	A	T	C	A	A	G
	G	G	A	A	C	C	G	A	T	G	C	T	A	A	G	G	T	A	A	A	A	T	T	G	A	T	A	A	A	A

Figure 2a(i)

SUBSTITUTE SHEET (RULE 26)

6/40

		250	260	270
<i>F.viral</i>		C A A G A A T T A G A T A A A T A T A A A A A T G C T G T A		
<i>F</i>		C A A G A A T T A G A T A A A T A T A A A A A T G C T G T A		
<i>F.opt</i>		C A A G A G C T G G A C A A G T A C A A G A A C G C C G T G		
		C A A G A A T T A G A T A A A T A T A A A A A T G C T G T A		
		280	290	300
<i>F.viral</i>		A C A G A A T T G C A G T T G C T C A T G C A A A G C A C A		
<i>F</i>		A C A G A A T T G C A G T T G C T C A T G C A G T C G A C A		
<i>F.opt</i>		A C C G A G C T G C A A C T G C T G A T G C A G T C G A C T		
		A C A G A A T T G C A G T T G C T C A T G C A G T C G A C A		
		310	320	330
<i>F.viral</i>		C A A G C A A C A A A C A A T C G A G C C A G A A G A G A A		
<i>F</i>		C A A G C A A C A A A C A A T C G A G C C A G A A G A G A A		
<i>F.opt</i>		C A A G C C A C C A A C A A C A G A G C C C G C A G A G A G		
		C A A G C A A C A A A C A A T C G A G C C A G A A G A G A A		
		340	350	360
<i>F.viral</i>		C T A C C A A G G T T T A T G A A T T A T A C A C T C A A C		
<i>F</i>		C T A C C T A G G T T T A T G A A T T A T A C A C T C A A C		
<i>F.opt</i>		C T G C C C C G C T T C A T G A A C T A C A C C C T G A A C		
		C T A C C A G G T T T A T G A A T T A T A C A C T C A A C		
		370	380	390
<i>F.viral</i>		A A T G C C A A A A A A A C C A A T G T A A C A T T A A G C		
<i>F</i>		A A T G C C A A A A A A A C C A A T G T A A C A C T T T C G		
<i>F.opt</i>		A A C G C C A A G A A G A C C A A C G T G A C C C T G T C C		
		A A T G C C A A A A A A A C C A A T G T A A C A C T T C C		
		400	410	420
<i>F.viral</i>		A A G A A A A G G A A A A G A A G A T T T C T T G G T T T T		
<i>F</i>		A A G A A A A G G A A A A G A A G A T T T C T T G G T T T T		
<i>F.opt</i>		A A G A A G A G G A A G C G C C G C T T C C T G G G C T T C		
		A A G A A A A G G A A A A G A A G A T T T C T T G G T T T T		
		430	440	450
<i>F.viral</i>		T T G T T A G G T G T T G G A T C T G C A A T C G C C A G T		
<i>F</i>		T T G T T A G G T G T T G G A T C C G C A A T C G C C A G T		
<i>F.opt</i>		C T G C T G G G C G T G G G C T C C G C C A T T G C C A G T		
		T T G T T A G G T G T T G G A T C C G C A A T C G C C A G T		
		460	470	480
<i>F.viral</i>		G G C G T T G C T G T A T C T A A G G T C C T G C A C C T A		
<i>F</i>		G G C G T T G C T G T A T C T A A G G T C C T G C A T C T A		
<i>F.opt</i>		G G C G T G G C C G T G T C C A A G G T G C T G C A C C T G		
		G G C G T T G C T G T A T C T A A G G T C C T G C A C C T A		
		490	500	510
<i>F.viral</i>		G A A G G G G A A G T G A A C A A G A T C A A A A G T G C T		
<i>F</i>		G A G G G G G A A G T G A A C A A G A T C A A A A G T G C T		
<i>F.opt</i>		G A G G G C G A G G T G A A C A A G A T C A A G A G T G C C		
		G A G G G G G A A G T G A A C A A G A T C A A A A G T G C T		

Figure 2a(ii)

SUBSTITUTE SHEET (RULE 26)

7/40

		520	530	540
<i>F.viral</i>		CTACTATCCACAAACAAGGCTGTAGTCAGC		
<i>F</i>		CTACTATCCACAAACAAGGCTGTAGTCAGC		
<i>F.opt</i>		CTGCTGTCCACTAACAAAGGCCGTGGTGAGC		
		CTACTATCCACAAACAAGGCTGTAGTCAGC		
		550	560	570
<i>F.viral</i>		TTATCAAATGGAGTTAGTGTCTTAACCAAGC		
<i>F</i>		TTATCAAATGGAGTTAGTGTCTTAACCAAGC		
<i>F.opt</i>		CTGAGCAACGGCGTGAGTGTGCTGACTAGC		
		TTATCAAATGGAGTTAGTGTCTTAACCAAGC		
		580	590	600
<i>F.viral</i>		AAAGTGTTAGACCTCAAAAACCTATATAGAT		
<i>F</i>		AAAGTGTTAGACCTCAAAAACCTATATAGAT		
<i>F.opt</i>		AAAGGTGCTGGACCTGAAGAACTACATCGAC		
		AAAGTGTTAGACCTCAAAAACCTATATAGAT		
		610	620	630
<i>F.viral</i>		AAACAATTGTTACCTATTGTGAACAAGCAA		
<i>F</i>		AAACAATTGTTACCTATTGTGAACAAGCAA		
<i>F.opt</i>		AAGCAATTGCTGCCCATTCGTGAACAAGCAG		
		AAACAATTGTTACCTATTGTGAACAAGCAA		
		640	650	660
<i>F.viral</i>		AGCTGCGAGCATATCAAATATAGAAACTGTG		
<i>F</i>		AGCTGCGAGCATATCAAATATAGAAACTGTG		
<i>F.opt</i>		TCCTGTAGCATCTCCAAACATCGAGACTGTG		
		AGCTGCGAGCATATCAAATATAGAAACTGTG		
		670	680	690
<i>F.viral</i>		ATAGAGTTCCAAACAAGAACAACAGACTA		
<i>F</i>		ATAGAGTTCCAAACAAGAACAACAGACTA		
<i>F.opt</i>		ATCGAGTTCCAGCAGAGAACAACCGGCCTG		
		ATAGAGTTCCAAACAAGAACAACAGACTA		
		700	710	720
<i>F.viral</i>		CTAGAGATTACCAGGGAATTTAGTGTTAAT		
<i>F</i>		CTAGAGATTACCAGGGAATTTAGTGTTAAT		
<i>F.opt</i>		CTGGAAATCACCCGGGAGTTCAQTGTGAAC		
		CTAGAGATTACCAGGGAATTTAGTGTTAAT		
		730	740	750
<i>F.viral</i>		GCAGGTGTAACCTACACCTGTAAAGCACTTAC		
<i>F</i>		GCAGGTGTAACCTACACCTGTAAAGCACTTAC		
<i>F.opt</i>		GCTGGCGTGACCACCTCCTGTCTCCACCTTAC		
		GCAGGTGTAACCTACACCTGTAAAGCACTTAC		
		760	770	780
<i>F.viral</i>		ATGTTAACTAATAGTGAATTATTGTCATTA		
<i>F</i>		ATGTTAACTAATAGTGAATTATTGTCATTA		
<i>F.opt</i>		ATGCTGACCAACAGCAGAGCTGCTGAGCCTG		
		ATGTTAACTAATAGTGAATTATTGTCATTA		

Figure 2a(iii)

SUBSTITUTE SHEET (RULE 26)

8/40

		790	800	810
<i>F.viral</i>		A T C A A T G A T A T G C C T A T A A C A A A T G A T C A G		
<i>F</i>		A T C A A T G A T A T G C C T A T A A C A A A T G A T C A G		
<i>F.opt</i>		A T C A A C G A C A T G C C C A T C A C C A A C G A C C A G		
		A T C A A T G A T A T G C C T A T A A C A A A T G A T C A G		
		820	830	840
<i>F.viral</i>		A A A A A G T T A A T G T C C A A C A A T G T T C A A A T A		
<i>F</i>		A A A A A G T T A A T G T C C A A C A A T G T T C A A A T A		
<i>F.opt</i>		A A G A A G C T T A T G T C C A A C A A C G T G C A G A T C		
		A A A A A G T T A A T G T C C A A C A A T G T T C A A A T A		
		850	860	870
<i>F.viral</i>		G T T A G A C A G C A A A G T T A C T C T A T C A T G T C C		
<i>F</i>		G T T A G A C A G C A A A G T T A C T C T A T C A T G T C C		
<i>F.opt</i>		G T G A G G C A G C A G A G C T A C T C C A T C A T G A G C		
		G T T A G A C A G C A A A G T T A C T C T A T C A T G T C C		
		880	890	900
<i>F.viral</i>		A T A A T A A A A G A G G A A G T C T T A G C A T A T G T A		
<i>F</i>		A T A A T A A A A G A G G A A G T C T T A G C A T A T G T A		
<i>F.opt</i>		A T C A T C A A G G A G G A G G T G C T G G C C T A T G T G		
		A T A A T A A A A G A G G A A G T C T T A G C A T A T G T A		
		910	920	930
<i>F.viral</i>		G T A C A A T T A C C A C T A T A T G G T G T T A T A G A T		
<i>F</i>		G T A C A A T T A C C A C T A T A T G G T G T T A T A G A T		
<i>F.opt</i>		G T G C A G C T G C C C C T G T A C G G C G T C A T C G A T		
		G T A C A A T T A C C A C T A T A T G G T G T T A T A G A T		
		940	950	960
<i>F.viral</i>		A C A C C C T G T T G G A A A C T A C A C A C A T C C C C T		
<i>F</i>		A C A C C C T G T T G G A A A C T A C A C A C A T C C C C T		
<i>F.opt</i>		A C C C C T T G C T G G A A G C T G C A C A C A G C C C C		
		A C A C C C T G T T G G A A A C T A C A C A C A T C C C C T		
		970	980	990
<i>F.viral</i>		C T A T G T A C A A C C A A C A C A A A A G A A G G G T C C		
<i>F</i>		C T A T G T A C A A C C A A C A C A A A A G A A G G G T C C		
<i>F.opt</i>		C T G T G C A C C A C C A A C A C A A A G G A G G G C A G C		
		C T A T G T A C A A C C A A C A C A A A A G A A G G G T C C		
		1000	1010	1020
<i>F.viral</i>		A A C A T C T G T T T A A C A A G A A A C T G A C A G A G G A		
<i>F</i>		A A C A T C T G T T T A A C A A G A A A C T G A C A G A G G A		
<i>F.opt</i>		A A C A T C T G C C T G A C C C G G A C C G A C C G C G G C		
		A A C A T C T G T T T A A C A A G A A A C T G A C A G A G G A		
		1030	1040	1050
<i>F.viral</i>		T G G T A C T G T G A C A A T G C A G G A T C A G T A T C T		
<i>F</i>		T G G T A C T G T G A C A A T G C A G G A T C A G T A T C T		
<i>F.opt</i>		T G G T A C T G T G A C A A C G C T G G C T C G G T G A G C		
		T G G T A C T G T G A C A A T G C A G G A T C A G T A T C T		

Figure 2a(iv)

SUBSTITUTE SHEET (RULE 26)

9/40

		1060	1070	1080
<i>F.viral</i>		T T C T T C C C A C A A G C T G A A A C A T G T A A A G T T		
<i>F</i>		T T C T T C C C A C A A G C T G A A A C A T G T A A A G T T		
<i>F.opt</i>		T T C T T C C C T C A A G C T G A A A C C T G C A A G G T C		
		T T C T T C C C A C A A G C T G A A A C A T G T A A A G T T		
		1090	1100	1110
<i>F.viral</i>		C A A T C A A A T C G A G T A T T T T G T G A C A C A A T G		
<i>F</i>		C A A T C A A A T C G A G T A T T T T G T G A C A C A A T G		
<i>F.opt</i>		C A G A G C A A C A G A G T G T T C T G T G A C A C C A T G		
		C A A T C A A A T C G A G T A T T T T G T G A C A C A A T G		
		1120	1130	1140
<i>F.viral</i>		A A C A G T T T A A C A T T A C C A A G T G A A G T A A A T		
<i>F</i>		A A C A G T T T A A C A T T A C C A A G T G A A G T A A A T		
<i>F.opt</i>		A A C T C C C T G A C C C T G C C C T C C G A G G T G A A C		
		A A C A G T T T A A C A T T A C C A A G T G A A G T A A A T		
		1150	1160	1170
<i>F.viral</i>		C T C T G C A A T G T T G A C A T A T T C A A C C C C A A A		
<i>F</i>		C T C T G C A A T G T T G A C A T A T T C A A C C C C A A A		
<i>F.opt</i>		C T G T G C A A C G T G G A T A T C T T C A A C C C C A A G		
		C T C T G C A A T G T T G A C A T A T T C A A C C C C A A A		
		1180	1190	1200
<i>F.viral</i>		T A T G A T T G T A A A A T T A T G A C T T C A A A A A C A		
<i>F</i>		T A T G A T T G T A A A A T T A T G A C T T C A A A A A C A		
<i>F.opt</i>		T A T G A C T G C A A G A T C A T G A C C T C C A A G A C C		
		T A T G A T T G T A A A A T T A T G A C T T C A A A A A C A		
		1210	1220	1230
<i>F.viral</i>		G A T G T A A G C A G C T C C G T T A T C A C A T C T C T A		
<i>F</i>		G A T G T A A G C A G C T C C G T T A T C A C A T C T C T A		
<i>F.opt</i>		G A T G T C T C G A G C T C C G T G A T C A C C A G C C T G		
		G A T G T A A G C A G C T C C G T T A T C A C A T C T C T A		
		1240	1250	1260
<i>F.viral</i>		G G A G C C A T T G T G T C A T G C T A T G G C A A A A C T		
<i>F</i>		G G A G C C A T T G T G T C A T G C T A T G G C A A A A C T		
<i>F.opt</i>		G G C G C C A T C G T G T C C T G C T A T G G C A A G A C C		
		G G A G C C A T T G T G T C A T G C T A T G G C A A A A C T		
		1270	1280	1290
<i>F.viral</i>		A A A T G T A C A G C A T C C A A T A A A A A T C G T G G A		
<i>F</i>		A A A T G T A C A G C A T C C A A T A A A A A T C G T G G A		
<i>F.opt</i>		A A G T G C A C C G C C A G C A A C A A G A A C C G G G G C		
		A A A T G T A C A G C A T C C A A T A A A A A T C G T G G A		
		1300	1310	1320
<i>F.viral</i>		A T C A T A A A G A C A T T T T C T A A C G G G T G C G A T		
<i>F</i>		A T C A T A A A G A C A T T T T C T A A C G G G T G C G A T		
<i>F.opt</i>		A T C A T C A A G A C C T T C A G C A A T G G G T G C G A C		
		A T C A T A A A G A C A T T T T C T A A C G G G T G C G A T		

Figure 2a(v)

SUBSTITUTE SHEET (RULE 26)

10/40

		1330	1340	1350
<i>F.viral</i>		T A T G T A T C A A A T A A A G G G G T G G A C A C T G T G		
<i>F</i>		T A T G T A T C A A A T A A A G G G G T G G A C A C T G T G		
<i>F.opt</i>		T A C G T T T C G A A C A A G G G C G T G G A C A C T G T G		
		T A T G T A T C A A A T A A A G G G G T G G A C A C T G T G		
		1360	1370	1380
<i>F.viral</i>		T C T G T A G G T A A C A C A T T A T A T T A T G T A A A T		
<i>F</i>		T C T G T A G G T A A C A C A T T A T A T T A T G T A A A T		
<i>F.opt</i>		T C C G T G G G C A A C A C C C T G T A C T A C G T G A A C		
		T C T G T A G G T A A C A C A T T A T A T T A T G T A A A T		
		1390	1400	1410
<i>F.viral</i>		A A G C A A G A A G G T A A A A G T C T C T A T G T A A A A		
<i>F</i>		A A G C A A G A A G G T A A A A G T C T C T A T G T A A A A		
<i>F.opt</i>		A A G C A A G A A G G G C A A G A G C C T G T A T G T G A A G		
		A A G C A A G A A G G T A A A A G T C T C T A T G T A A A A		
		1420	1430	1440
<i>F.viral</i>		G G T G A A C C A A T A A T A A A T T T C T A T G A C C C A		
<i>F</i>		G G T G A A C C A A T A A T A A A T T T C T A T G A C C C A		
<i>F.opt</i>		G G C G A G C C C A T C A T C A A C T T C T A C G A C C C C		
		G G T G A A C C A A T A A T A A A T T T C T A T G A C C C A		
		1450	1460	1470
<i>F.viral</i>		T T A G T A T T C C C C T C T G A T G A A T T T G A T G C A		
<i>F</i>		T T A G T A T T C C C C T C T G A T G A A T T T G A T G C G		
<i>F.opt</i>		C T G G T G T T C C C C T C C G A C G A A T T C G A C G C C		
		T T A G T A T T C C C C T C T G A T G A A T T T G A C G C		
		1480	1490	1500
<i>F.viral</i>		T C A A T A T C T C A A G T C A A C G A G A A G A T T A A C		
<i>F</i>		T C A A T A T C T C A A G T C A A C G A G A A G A T T A A C		
<i>F.opt</i>		T C C A T T A G C C A A G T C A A C G A G A A G A T C A A C		
		T C A A T A T C T C A A G T C A A C G A G A A G A T T A A C		
		1510	1520	1530
<i>F.viral</i>		C A G A G C C T A G C A T T T A T T C G T A A A T C C G A T		
<i>F</i>		C A G A G C C T A G C A T T T A T T C G T A A A T C C G A T		
<i>F.opt</i>		C A G A G C C T G G C C T T C A T C C G C A A G T C C G A C		
		C A G A G C C T A G C A T T T A T T C G T A A A T C C G A T		
		1540	1550	1560
<i>F.viral</i>		G A A T T A T T A C A T A A T G T A A A T G C T G G T A A A		
<i>F</i>		G A A T T A T T A C A T A A T G T A A A T G C T G G G A A G		
<i>F.opt</i>		G A G C T G C T G C A C A A C G T C A A C G C T G G C A A G		
		G A A T T A T T A C A T A A T G T A A A T G C T G G A A G		
		1570	1580	1590
<i>F.viral</i>		T C C A C C A C A A A T A T C A T G A T A A C T A C T A T A		
<i>F</i>		A G C A C C A C A A A T A T C A T G A T A A C T A C T A T A		
<i>F.opt</i>		A G C A C C A C C A A C A T C A T G A T C A C C A C C A T C		
		A G C A C C A C A A A T A T C A T G A T A A C T A C T A T A		

Figure 2a(vi)

SUBSTITUTE SHEET (RULE 26)

11/40

		1600	1610	1620
<i>F.viral</i>		A T T A T A G T G A T T A T A G T A A T A T T G T T A T C A		
<i>F</i>		A T T A T A G T G A T T A T A G T A A T A T T G T T A T C A		
<i>F.opt</i>		A T C A T C G T G A T C A T C G T G A T C C T G C T G A G C		
		A T T A T A G T G A T T A T A G T A A T A T T G T T A T C A		
		1630	1640	1650
<i>F.viral</i>		T T A A T T G C T G T T G G A C T G C T C T T A T A C T G T		
<i>F</i>		T T A A T T G C T G T T G G A C T G C T C T T A T A C T G T		
<i>F.opt</i>		C T G A T C G C C G T G G G C C T G C T G C T G T A C T G C		
		T T A A T T G C T G T T G G A C T G C T C T T A T A C T G T		
		1660	1670	1680
<i>F.viral</i>		A A G G C C A G A A G C A C A C C A G T C A C A C T A A G C		
<i>F</i>		A A G G C C A G A T C T A C A C C A G T C A C A C T A A G C		
<i>F.opt</i>		A A G G C C C G G A G C A C T C C C G T G A C C O T G A G C		
		A A G G C C A G A A G C A C A C C A G T C A C A C T A A G C		
		1690	1700	1710
<i>F.viral</i>		A A A G A T C A A C T G A G T G G T A T A A A T A A T A T T		
<i>F</i>		A A A G A T C A A C T G A G T G G T A T A A A T A A T A T T		
<i>F.opt</i>		A A G G A C C A G C T G A G C G G C A T C A A C A A C A T C		
		A A A G A T C A A C T G A G T G G T A T A A A T A A T A T T		
		1720	1730	1740
<i>F.viral</i>		G C A T T T A G T A A C T A A		
<i>F</i>		G C A T T T A G T A A C T A A		
<i>F.opt</i>		G C C T T C A G C A A C T G A		
		G C A T T T A G T A A C T A A		

Figure 2a(vii)

SUBSTITUTE SHEET (RULE 26)

12/40

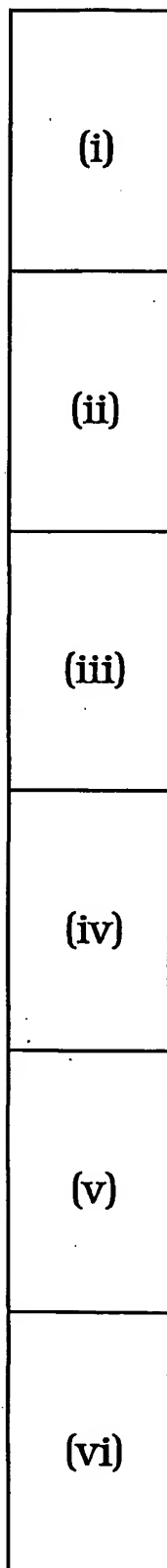


Figure 2b
SUBSTITUTE SHEET (RULE 26)

ClustalW Formatted Alignments

SUBSTITUTE SHEET (RULE 26)

14/40

	250	260	270
<i>F.sol.viral</i>	C A A G A A T T A G A T A A A T A T A A A A A T G C T G T A		
<i>F.sol</i>	C A A G A A T T A G A T A A A T A T A A A A A T G C T G T A		
<i>F.sol.opt</i>	C A A G A G C T G G A C A A G T A C A A G A A C G C C G T G		
	C A A G A A T T A G A T A A A T A T A A A A A T G C T G T A		
	280	290	300
<i>F.sol.viral</i>	A C A G A A T T G C A G T T G C T C A T G C A A A G C A C A		
<i>F.sol</i>	A C A G A A T T G C A G T T G C T C A T G C A G T C G A C A		
<i>F.sol.opt</i>	A C C G A G C T G C A A C T G C T G A T G C A G T C G A C T		
	A C A G A A T T G C A G T T G C T C A T G C A G T C G A C A		
	310	320	330
<i>F.sol.viral</i>	C A A G C A A C A A A C A A T C G A G C C A G A A G A G A A		
<i>F.sol</i>	C A A G C A A C A A A C A A T C G A G C C A G A A G A G A A		
<i>F.sol.opt</i>	C A A G C C A C C A A C A A C A G A G C C C G C A G A G A G		
	C A A G C A A C A A A C A A T C G A G C C A G A A G A G A A		
	340	350	360
<i>F.sol.viral</i>	C T A C C A A G G T T T A T G A A T T A T A C A C T C A A C		
<i>F.sol</i>	C T A C C T A G G T T T A T G A A T T A T A C A C T C A A C		
<i>F.sol.opt</i>	C T G C C C C G C T T C A T G A A C T A C A C C C T G A A C		
	C T A C C A G G T T T A T G A A T T A T A C A C T C A A C		
	370	380	390
<i>F.sol.viral</i>	A A T G C C A A A A A A A C C A A T G T A A C A T T A A G C		
<i>F.sol</i>	A A T G C C A A A A A A A C C A A T G T A A C A C T T T C G		
<i>F.sol.opt</i>	A A C G C C A A G A A G A C C A A C G T G A C C C T G T C C		
	A A T G C C A A A A A A A C C A A T G T A A C A C T T C C		
	400	410	420
<i>F.sol.viral</i>	A A G A A A A G G A A A A G A A G A T T T C T T G G T T T T		
<i>F.sol</i>	A A G A A A A G G A A A A G A A G A T T T C T T G G T T T T		
<i>F.sol.opt</i>	A A G A A G A G G A A G C G C C G C T T C C T G G G C T T C		
	A A G A A A A G G A A A A G A A G A T T T C T T G G T T T T		
	430	440	450
<i>F.sol.viral</i>	T T G T T A G G T G T T G G A T C T G C A A T C G C C A G T		
<i>F.sol</i>	T T G T T A G G T G T T G G A T C C G C A A T C G C C A G T		
<i>F.sol.opt</i>	C T G C T G G G C G T G G G C T C C G C C A T T G C C A G T		
	T T G T T A G G T G T T G G A T C C G C A A T C G C C A G T		
	460	470	480
<i>F.sol.viral</i>	G G C G T T G C T G T A T C T A A G G T C C T G C A C C T A		
<i>F.sol</i>	G G C G T T G C T G T A T C T A A G G T C C T G C A T C T A		
<i>F.sol.opt</i>	G G C G T G G C C G T G T C C A A G G T G C T G C A C C T G		
	G G C G T T G C T G T A T C T A A G G T C C T G C A C C T A		
	490	500	510
<i>F.sol.viral</i>	G A A G G G G A A G T G A A C A A G A T C A A A A G T G C T		
<i>F.sol</i>	G A G G G G G A A G T G A A C A A G A T C A A A A G T G C T		
<i>F.sol.opt</i>	G A G G G G C G A G G T G A A C A A G A T C A A G A G T G C C		
	G A G G G G G A A G T G A A C A A G A T C A A A A G T G C T		

Figure 2b(ii)

SUBSTITUTE SHEET (RULE 26)

15/40

	520	530	540
<i>F.sol.viral</i>	CTACTATCCACAACAAGGCTGTAGTCAGC		
<i>F.sol</i>	CTACTATCCACAACAAGGCTGTAGTCAGC		
<i>F.sol.opt</i>	CTGCTGTCCACTAACAAAGGCCCGTGGTGAGC		
	CTACTATCCACAACAAGGCTGTAGTCAGC		
	550	560	570
<i>F.sol.viral</i>	TTATCAAAATGGAGTTAGTGTCTTAACCAGC		
<i>F.sol</i>	TTATCAAAATGGAGTTAGTGTCTTAACCAGC		
<i>F.sol.opt</i>	CTGAGCAACGGCGTGAGTGTGCTGACTAGC		
	TTATCAAAATGGAGTTAGTGTCTTAACCAGC		
	580	590	600
<i>F.sol.viral</i>	AAAGTGTTAGACCTCAAAAACCTATATAGAT		
<i>F.sol</i>	AAAGTGTTAGACCTCAAAAACCTATATAGAT		
<i>F.sol.opt</i>	AAAGGTGCTGGACCTGAAGAACTACATCGAC		
	AAAGTGTTAGACCTCAAAAACCTATATAGAT		
	610	620	630
<i>F.sol.viral</i>	AAACAATTGTTACCTATTGTGAACAAGCAA		
<i>F.sol</i>	AAACAATTGTTACCTATTGTGAACAAGCAA		
<i>F.sol.opt</i>	AAAGCAATTGCTGGCCCATCGTGAACAAGCAG		
	AAACAATTGTTACCTATTGTGAACAAGCAA		
	640	650	660
<i>F.sol.viral</i>	AGCTGCGAGCATATCAAAATATAGAAACTGTG		
<i>F.sol</i>	AGCTGCGAGCATATCAAAATATAGAAACTGTG		
<i>F.sol.opt</i>	TCCCTGTAGCATCTCCAAACATCGAGACTGTG		
	AGCTGCGAGCATATCAAAATATAGAAACTGTG		
	670	680	690
<i>F.sol.viral</i>	ATAGAGTTTCCAACAAGAGAACCAACAGACTA		
<i>F.sol</i>	ATAGAGTTTCCAACAAGAGAACCAACAGACTA		
<i>F.sol.opt</i>	ATCGAGTTTCCAGCAGAAAGAACCAACCGCCTG		
	ATAGAGTTTCCAACAAGAGAACCAACAGACTA		
	700	710	720
<i>F.sol.viral</i>	CTAGAGATTACCAGGGAATTTAGTGTTAAT		
<i>F.sol</i>	CTAGAGATTACCAGGGAATTTAGTGTTAAT		
<i>F.sol.opt</i>	CTGGAAATCACCCGGGAGTTTCAGTGTGAAC		
	CTAGAGATTACCAGGGAATTTAGTGTTAAT		
	730	740	750
<i>F.sol.viral</i>	GCAGGTGTAACCTACACCTGTAAAGCACTTAC		
<i>F.sol</i>	GCAGGTGTAACCTACACCTGTAAAGCACTTAC		
<i>F.sol.opt</i>	GCTGGCGTGACCACTCCTGTCTCCACCTTAC		
	GCAGGTGTAACCTACACCTGTAAAGCACTTAC		
	760	770	780
<i>F.sol.viral</i>	ATGTTAACTAATAGTGAATTATTGTCATTAA		
<i>F.sol</i>	ATGTTAACTAATAGTGAATTATTGTCATTAA		
<i>F.sol.opt</i>	ATGCTGACCAACAGCGAGCTGCTGAGCCTG		
	ATGTTAACTAATAGTGAATTATTGTCATTAA		

Figure 2b(iii)

SUBSTITUTE SHEET (RULE 26)

16/40

		790	800	810
<i>F.sol.viral</i>		ATCAATGATATGCCTATAACAAATGATCAG		
<i>F.sol</i>		ATCAATGATATGCCTATAACAAATGATCAG		
<i>F.sol.opt</i>		ATCAACGACATGCCCATCACCAACGACCCAG		
		ATCAATGATATGCCTATAACAAATGATCAG		
		820	830	840
<i>F.sol.viral</i>		AAAAAGTTAATGTCCAACAATGTTCAAATA		
<i>F.sol</i>		AAAAAGTTAATGTCCAACAATGTTCAAATA		
<i>F.sol.opt</i>		AAGAAGCTTATGTCCAACAACGTTGCAGATC		
		AAAAAGTTAATGTCCAACAATGTTCAAATA		
		850	860	870
<i>F.sol.viral</i>		GTTAGACAGCAAAGTTACTCTATCATGTTCC		
<i>F.sol</i>		GTTAGACAGCAAAGTTACTCTATCATGTTCC		
<i>F.sol.opt</i>		GTGAGGCAGCAGAGCTACTCCATCATGAGC		
		GTTAGACAGCAAAGTTACTCTATCATGTTCC		
		880	890	900
<i>F.sol.viral</i>		ATAATAAAGAGGGAAGTCTTAGCATATGTA		
<i>F.sol</i>		ATAATAAAGAGGGAAGTCTTAGCATATGTA		
<i>F.sol.opt</i>		ATCATCAAGGAGGAGGTGCTGGCCTATGTG		
		ATAATAAAGAGGGAAGTCTTAGCATATGTA		
		910	920	930
<i>F.sol.viral</i>		GTACAATTACCACTATATGGTGTTATAGAT		
<i>F.sol</i>		GTACAATTACCACTATATGGTGTTATAGAT		
<i>F.sol.opt</i>		GTGCAGCTGCCCTGTACGGCGTCAATCGAT		
		GTACAATTACCACTATATGGTGTTATAGAT		
		940	950	960
<i>F.sol.viral</i>		ACACCCTGTGTGGAAACTACACACATCCCT		
<i>F.sol</i>		ACACCCTGTGTGGAAACTACACACATCCCT		
<i>F.sol.opt</i>		ACCCCTTGTCTGGAAAGCTGCACACAGCCCT		
		ACACCCTGTGTGGAAACTACACACATCCCT		
		970	980	990
<i>F.sol.viral</i>		CTATGTACAACCAACACAAAAGAAAGGGTCC		
<i>F.sol</i>		CTATGTACAACCAACACAAAAGAAAGGGTCC		
<i>F.sol.opt</i>		CTGTGCACCAACCAACACCAAGGAAGGGCAGC		
		CTATGTACAACCAACACAAAAGAAAGGGTCC		
		1000	1010	1020
<i>F.sol.viral</i>		AACATCTGTTTAACAAGAACTGACAGAGGA		
<i>F.sol</i>		AACATCTGTTTAACAAGAACTGACAGAGGA		
<i>F.sol.opt</i>		AACATCTGCTTGACCCGGACCGACCGCGGC		
		AACATCTGTTTAACAAGAACTGACAGAGGA		
		1030	1040	1050
<i>F.sol.viral</i>		TGGTACTGTGACAATGCAGGATCAGTATCT		
<i>F.sol</i>		TGGTACTGTGACAATGCAGGATCAGTATCT		
<i>F.sol.opt</i>		TGGTACTGTGACAACGCTGGCTCGGTGAGC		
		TGGTACTGTGACAATGCAGGATCAGTATCT		

Figure 2b(iv)

SUBSTITUTE SHEET (RULE 26)

17/40

		1060	1070	1080
<i>F.sol.viral</i>		T T C T T C C C A C A A G C T G A A A C A T G T A A A G T T		
<i>F.sol</i>		T T C T T C C C A C A A G C T G A A A C A T G T A A A G T T		
<i>F.sol.opt</i>		T T C T T C C C T C A A G C T G A A A C C T G C A A G G T C		
		T T C T T C C C A C A A G C T G A A A C A T G T A A A G T T		
		1090	1100	1110
<i>F.sol.viral</i>		C A A T C A A A T C G A G T A T T T T G T G A C A C A A T G		
<i>F.sol</i>		C A A T C A A A T C G A G T A T T T T G T G A C A C A A T G		
<i>F.sol.opt</i>		C A G A G C A A C A G A G T G T T C T G T G A C A C C A T G		
		C A A T C A A A T C G A G T A T T T T G T G A C A C A A T G		
		1120	1130	1140
<i>F.sol.viral</i>		A A C A G T T T A A C A T T A C C A A G T G A A G T A A A T		
<i>F.sol</i>		A A C A G T T T A A C A T T A C C A A G T G A A G T A A A T		
<i>F.sol.opt</i>		A A C T C C C T G A C C C T G C C C T C C G A G G T G A A C		
		A A C A G T T T A A C A T T A C C A A G T G A A G T A A A T		
		1150	1160	1170
<i>F.sol.viral</i>		C T C T G C A A T G T T G A C A T A T T C A A C C C C A A A		
<i>F.sol</i>		C T C T G C A A T G T T G A C A T A T T C A A C C C C A A A		
<i>F.sol.opt</i>		C T G T G C A A C G T G G A T A T C T T C A A C C C C A A G		
		C T C T G C A A T G T T G A C A T A T T C A A C C C C A A A		
		1180	1190	1200
<i>F.sol.viral</i>		T A T G A T T G T A A A A T T A T G A C T T C A A A A A C A		
<i>F.sol</i>		T A T G A T T G T A A A A T T A T G A C T T C A A A A A C A		
<i>F.sol.opt</i>		T A T G A C T G C A A G A T C A T G A C C T C C A A G A C C		
		T A T G A T T G T A A A A T T A T G A C T T C A A A A A C A		
		1210	1220	1230
<i>F.sol.viral</i>		G A T G T A A G C A G C T C C G T T A T C A C A T C T C T A		
<i>F.sol</i>		G A T G T A A G C A G C T C C G T T A T C A C A T C T C T A		
<i>F.sol.opt</i>		G A T G T C T C G A G C T C C G T G A T C A C C A G C C T G		
		G A T G T A A G C A G C T C C G T T A T C A C A T C T C T A		
		1240	1250	1260
<i>F.sol.viral</i>		G G A G C C A T T G T G T C A T G C T A T G G C A A A A C T		
<i>F.sol</i>		G G A G C C A T T G T G T C A T G C T A T G G C A A A A C T		
<i>F.sol.opt</i>		G G C G C C A T C G T G T C C T G C T A T G G C A A G A C C		
		G G A G C C A T T G T G T C A T G C T A T G G C A A A A C T		
		1270	1280	1290
<i>F.sol.viral</i>		A A A T G T A C A G C A T C C A A T A A A A A T C G T G G A		
<i>F.sol</i>		A A A T G T A C A G C A T C C A A T A A A A A T C G T G G A		
<i>F.sol.opt</i>		A A G T G C A C C G C C A G C A A C A A G A A C C G G G G C		
		A A A T G T A C A G C A T C C A A T A A A A A T C G T G G A		
		1300	1310	1320
<i>F.sol.viral</i>		A T C A T A A A G A C A T T T T C T A A C G G G T G C G A T		
<i>F.sol</i>		A T C A T A A A G A C A T T T T C T A A C G G G T G C G A T		
<i>F.sol.opt</i>		A T C A T C A A G A C C T T T C A G C A A T G G G T G C G A C		
		A T C A T A A A G A C A T T T T C T A A C G G G T G C G A T		

Figure 2b(v)

SUBSTITUTE SHEET (RULE 26)

18/40

	1330	1340	1350
<i>F.sol.viral</i>	T A T G T A T C A A A T A A A G G G G T G G A C A C T G T G		
<i>F.sol</i>	T A T G T A T C A A A T A A A G G G G T G G A C A C T G T G		
<i>F.sol.opt</i>	T A C G T T T C G A A C A A G G G C G T G G A C A C T G T G		
	T A T G T A T C A A A T A A A G G G G T G G A C A C T G T G		
	1360	1370	1380
<i>F.sol.viral</i>	T C T G T A G G T A A C A C A T T A T A T T A T G T A A A T		
<i>F.sol</i>	T C T G T A G G T A A C A C A T T A T A T T A T G T A A A T		
<i>F.sol.opt</i>	T C C G T G G G C A A C A C C C T G T A C T A C G T G A A C		
	T C T G T A G G T A A C A C A T T A T A T T A T G T A A A T		
	1390	1400	1410
<i>F.sol.viral</i>	A A G C A A G A A G G T A A A A G T C T C T A T G T A A A A		
<i>F.sol</i>	A A G C A A G A A G G T A A A A G T C T C T A T G T A A A A		
<i>F.sol.opt</i>	A A G C A A G A A G G G C A A G A G C C T G T A T G T G A A G		
	A A G C A A G A A G G T A A A A G T C T C T A T G T A A A A		
	1420	1430	1440
<i>F.sol.viral</i>	G G T G A A C C A A T A A T A A A T T T C T A T G A C C C A		
<i>F.sol</i>	G G T G A A C C A A T A A T A A A T T T C T A T G A C C C A		
<i>F.sol.opt</i>	G G C G A G C C C A T C A T C A A C T T C T A C G A C C C C		
	G G T G A A C C A A T A A T A A A T T T C T A T G A C C C A		
	1450	1460	1470
<i>F.sol.viral</i>	T T A G T A T T C C C C T C T G A T G A A T T T G A T G C A		
<i>F.sol</i>	T T A G T A T T C C C C T C T G A T G A A T T T G A C G C G		
<i>F.sol.opt</i>	C T G G T G T T C C C C T C C G A C G A A T T C G A C G C C		
	T T A G T A T T C C C C T C T G A T G A A T T T G A C G C		
	1480	1490	1500
<i>F.sol.viral</i>	T C A A T A T C T C A A G T C A A C G A G A A G A T T A A C		
<i>F.sol</i>	T C A A T A T C T C A A G T C A A C G A G A A G A T T A A C		
<i>F.sol.opt</i>	T C C A T T A G C C A A G T C A A C G A G A A G A T C A A C		
	T C A A T A T C T C A A G T C A A C G A G A A G A T T A A C		
	1510	1520	1530
<i>F.sol.viral</i>	C A G A G C C T A G C A T T T A T T C G T A A A T C C G A T		
<i>F.sol</i>	C A G A G C C T A G C A T T T A T T C G T A A A T C C G A T		
<i>F.sol.opt</i>	C A G A G C C T G G C C T T C A T C C G C A A G T C C G A C		
	C A G A G C C T A G C A T T T A T T C G T A A A T C C G A T		
	1540	1550	1560
<i>F.sol.viral</i>	G A A T T A T T A C A T A A T G T A A A T G C T G G T A A A		
<i>F.sol</i>	G A A T T A T T A C A T A A T G T A A A T G C T G G G A A G		
<i>F.sol.opt</i>	G A G C T G C T G C A C A A C G T C A A C G C T G G C A A G		
	G A A T T A T T A C A T A A T G T A A A T G C T G G A A G		
	1570	1580	1590
<i>F.sol.viral</i>	T C C A C C A C A A A T T A A		
<i>F.sol</i>	A G C A C C A C A A A T T A A		
<i>F.sol.opt</i>	A G C A C C A C C A A C T G A		
	A G C A C C A C A A A T T A A		

Figure 2b(vi)

SUBSTITUTE SHEET (RULE 26)

19/40

(i)
(ii)
(iii)
(iv)

Figure 3a
SUBSTITUTE SHEET (RULE 26)

410	420	430	440	450	460	470	480	490	500
CGTGACCTG	TCCAAGAAGA	GGAGGCGCG	CTTCTGGGC	TTCTGCTGG	CGGTGGGCTC	CGCCATTGCC	AGTGGCGTGG	CCGTGTCCAA	GGTGCTGCAC
GCACCTGGGAC	AGGTTCTTCT	CCTTCGGGC	GAAGGACCCG	AAGGACGACC	CGCACCCGAG	CGGTTAACGG	TCACCGCACC	GGCACAGGTT	CCACGACGTG
V T L S K K	R K R R	F L G F L L	G V G S A I A	S G V A V S K	V L H				

Figure 3a(i)

21/40

510	520	530	540	550	560	570	580	590	600
CTGGAGGGCG	AGGTGAACAA	GATCAAGAGT	GCCCTGCTGT	CCACTAACAA	GGCCCTGGTG	AGCCTGAGCA	ACGGCGTGAG	TGTGCTGACT	AGCAAGGTGC
GACCTCCCGC	TCCACTTGT	CTAGTTCTCA	CGGGACGACA	GGTGATTGTT	CCGGACCAAC	TCGGACTCGT	TGCCGCACTC	ACACGACTGA	TCGTTCCACG
L E G	E V N K	I K S	A L L	S T N K	A V V	S L S	N G V S	V L T	S K V
a	a	a	a	a	a	a	a	a	a
>MfeI									
610	620	630	640	650	660	670	680	690	700
TGGACCTGAA	GAACACATC	GACAAGCAAT	TGCTGCCCAT	CGTGAACAAG	CAGTCCTGTA	GCATCTCCAA	CATCGAGACT	GTGATCGAGT	TCCAGCAGAA
ACCTGGACTT	CTTGATGTAG	CTGTTCTGTTA	ACGACGGGTA	GCACTTGTTT	GTGAGGACT	GTAGCTCTGA	CACTAGCTCA	AGGTCGTCTT	
L D L K	N Y I	D K Q	L L P	I V N K	Q S C	S I S N	I E T	V I E	F Q Q K
a	a	a	a	a	a	a	a	a	a
>SmaI									
710	720	730	740	750	760	770	780	790	800
GAACAACCGC	CTGCTGGAAA	TCACCCGGGA	GTTCAAGTGT	AACGCTGGCG	TGACCACTCC	TGTCTCCACC	TACATGCTGA	CCAACAGCGA	GCTGCTGAGC
CTTGTGGCG	GACGACCTTT	AGTGGGCCCT	CAAGTCACAC	TTGCGACCGC	ACTGGTGAGG	ACAGAGGTGG	ATGTACGACT	GGTTGTGCGT	CGACGACTCG
N N R	L L E	I T R	E F S	V N A	G V T	T P V	S T Y	M L T	N S E L L S
a	a	a	a	a	a	a	a	a	a
>HindIII									
810	820	830	840	850	860	870	880	890	900
CTGATCAACG	ACATGCCCAT	CACCAACGAC	CAGAAGAAGC	TTATGTCCAA	CAACGTGCAG	ATCGTGAGGC	ACCAGAGCTA	CTCCATCATG	AGCATCATCA
GACTAGTTGC	TGTACGGGTA	GTGGTTGCTG	GTCTTCTTCG	AATACAGGTT	GTTGCACTGC	TAGCACTCCG	TCGTCTCGAT	GAGGTAGTAC	TCGTAGTAGT
L I N	D M P	I T N	D Q K	K L M	S N V	Q I V	R Q S	Y S I	M S I I
a	a	a	a	a	a	a	a	a	a
>ClaI									
910	920	930	940	950	960	970	980	990	1000
AGGAGGAGGT	GCTGGCCTAT	GTGGTGCAGC	TGCCCTGTGA	CGGCGTCAAC	GATACCCCTT	GCTGGAAGCT	GCACACCAGC	CCCCGTGTGA	CCACCAACAC
TCCTCCTCCA	CGACCGGATA	CACCAAGTGC	ACGGGGACAT	GCCGAGTAG	CTATGGGAA	CGACCTTCCA	CGTGTGGTGC	GGGACACGT	GGTGGTTGTG
K E E	V L A	Y V V	Q L P	L Y G	V I D	T P C	W K L	H T S	P L C T T N T
a	a	a	a	a	a	a	a	a	a

Figure 3a(ii)

22/40

>RsrII											
1010	1020	1030	1040	1050	1060	1070	1080	1090	1100		
CAAGGAGGC	AGCAACATCT	GCCTGACCCG	GACCGACCGC	GGCTGGTACT	GTGACACCGC	TGGCTCGGTG	AGCTTCTTCC	CTCAAGCTGA	AACCTGCAAG		
GTTCCTCCCG	TCGTTGTAGA	CGGACTGGGC	CTGGCTGGCG	CCGACCATGA	CACCTGTTGG	ACCGAGCCAC	TCGAAGAAGG	GAGTTCGACT	TTGGACGTTT		
K E G	S N I	C L T	R T	D R	G W Y	C D N	A G S	V S F	F P Q	A E T	C K
a	a	a	a	a	a	a	a	a	a	a	a
>EcoRV											
1110	1120	1130	1140	1150	1160	1170	1180	1190	1200		
GTCCAGAGCA	ACAGAGTGT	CTGTGACACC	ATGAACCTCC	TGACCCTGCC	CTCCGAGGTG	AACCTGTGCA	ACGTGGATAT	CTTCAACCCC	AAGTATGACT		
CAGGTCTCGT	TGTCTCACAA	GACACTGTGG	TACTTGAGGG	ACTGGGACGG	GAGGCTCCAC	TTGGACACGT	TGCACCTATA	GAAGTTGGGG	TTCATACTGA		
V Q S	N R V	F C D	T M N	S L T	P S E	V N L	C N V	D I F	N P K	Y D	
a	a	a	a	a	a	a	a	a	a	a	a
>XhoI											
1210	1220	1230	1240	1250	1260	1270	1280	1290	1300		
GCAAGATCAT	GACCTCCAAG	ACCGATGTCT	CGAGCTCCGT	GATCACCAGC	CTGGGCGCCA	TCGTGTCCTG	CTATGGCAAG	ACCAAGTGCA	CCGCCAGCAA		
CGTTCTAGTA	CTGGAGGTTT	TGGCTACAGA	GCTCGAGGCA	CTAGTGGTCG	GACCCGCGGT	AGCACAGGAC	GATACCGTTC	TGGTTCACGT	GGCGGTCGTT		
C K I	M T S	K T D	V S S	S V I	T S L	G A I	V S C	Y G K	T K C	T A S	N
a	a	a	a	a	a	a	a	a	a	a	a
>BstBI											
1310	1320	1330	1340	1350	1360	1370	1380	1390	1400		
CAAGAACCGG	GGCATCATCA	AGACCTTCAG	CAATGGGTGC	GACTACGTTT	CGAACACAGG	CGTGGACACT	GTGTCCGTGG	GCAACACCCCT	GTACTACGTG		
GTTCCTGGCC	CCGTAGTAGT	TCTGGAAGTC	GTTACCCACG	CTGATGCAAA	GCTTGTTCCC	GCACCTGTGA	CACAGGCACC	CGTTGTGGGA	CATGATGCAC		
K N R	G I I	K T F	S N G	C D Y	V S N	K G V	D T V	S V G	N T L	Y Y	V
a	a	a	a	a	a	a	a	a	a	a	a
>EcoRV											
1410	1420	1430	1440	1450	1460	1470	1480	1490	1500		
AACAAGCAAG	AGGGCAAGAG	CCGTATGTG	AAGGGCGAGC	CCATCATCAA	CTTCTACGAC	CCCCTGGTGT	TCCCTCCGA	CGAATTCGAC	GCCTCCATTA		
TTGTTCGTTT	TCCCGTTCTC	GGACATACAC	TTCCCGCTCG	GGTAGTAGTT	GAAGATGCTG	GGGGACCACA	AGGGGAGGCT	GCTTAAGCTG	CGGAGGTAAT		
N K Q	E G K	S L Y	V K G	E P I	I N F	Y D P	L V F	P S D	E F D	A S I	
a	a	a	a	a	a	a	a	a	a	a	a

Figure 3a(iii)

24/40

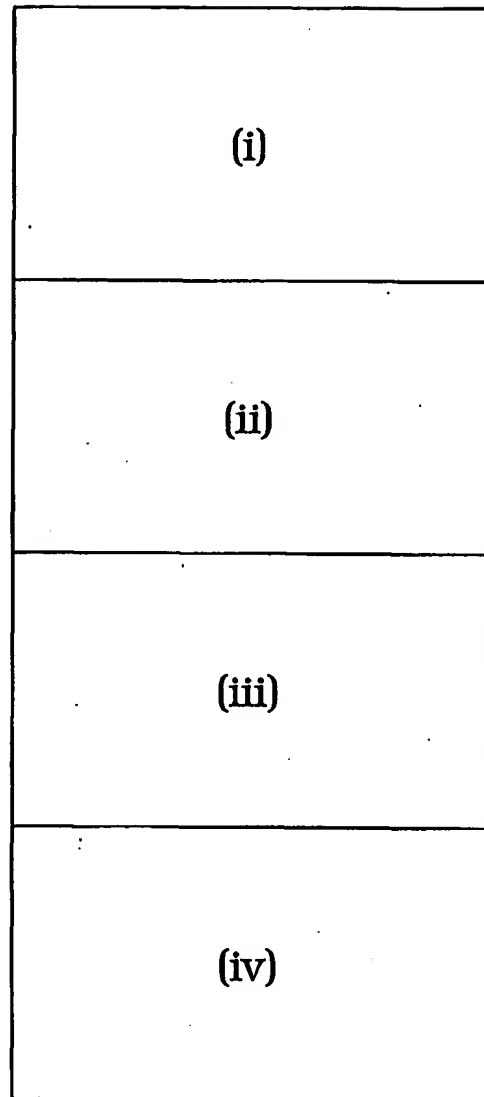


Figure 3b

>PstI
|
| 10 20 30 40 50 60 70 | 80 90 100
CTGCAGTCAC CGTCCCTTGAC ACCATGAGGC TGCTGATCCT GAAGGCCAAC GCCATCACCA CCATCCTGAC CGCGGTGACC TTCTGCTTCG CCTCTGGCCA
GACGCTCAGTG GCAGGAAGCTG TGGTACTCG ACGACTAGGA CTTCCGGTTG CGGTACTGGT GGTAGGACTG GCGCCACTGG AAGACAAGC GGAGACCGGT
M E L L I L K A N A I T T I L T A V T F C F A S G Q>
a a a a a a a a a a

>BstEII
|
| 110 120 130 140 150 160 170 | 180 190 200
GAACATCACT GAGGAGTTCT ACCAGAGCAC TTGTTCCGCT GTGAGCAAGG GCTACCTGAG CGCCCTGAGG ACCGGTTGGT ACACAGCGGT GATCACCATC
CTTGTAAGTA CTCCTCAAGA TGGTCTCGTG AACAAAGGCGA CACTCGTTCC CGATGACTC GCGGACTCC TGGCCAACCA TGTGTCGCA CTAGTGGTAG
N I T E E F Y Q S T C S A V S K G Y L S A L R T G W Y T S V I T I>
a a a a a a a a a a

>AgeI
|
| 210 220 230 240 250 260 270 | 280 290 300
GAGCTGAGCA ACATCAAGAA GAACAGTGC AACGGCACCG AGCCCAAGGT GAAGCTGATC AAGCAAGAGC TGGACAAGTA CAAGAAGGCC GTGRCCGAGC
CTCGACTCGT TGTAGTTCTT CTTGTTACAG TTGCCGTGGC TGGCGTTCCA CTTGACTAG TTCGTTCTCG ACCTGTTCTAT GTTCTGCGG CACTGGCTCG
E L S N I K K N K C N G T D A K V K L I K Q E L D K Y K N A V T E>
a a a a a a a a a a

>BstXI
|
| 310 320 330 340 350 360 370 | 380 390 400
TGCAACTGCT GATGCAGTCG ACTCAGGCA CCAACAACAG AGCCCCGAGA GAGCTGCCCC GCTTCATGAA CTACACCTCG AACAAAGCCA AGAAGACCAA
ACGTTGACGA CTACGTCAGC TGAGTTCGGT GGTGTTGTC TCGGGCGTCT CTCGACGGGG CGAAGTACTT GATGTGGGAC TTGTGCGGT TCTTCTGGTT
L Q L L M Q S T Q A T N N R A R R A R R E L P R F M N Y T L N N A K K T N>
a a a a a a a a a a

>BstXI
|
| 410 420 430 440 450 460 470 | 480 490 500
CGTGACCCCTG TCCAAGAAGA GGAAGCGCCG CTTCTGSGC TTCTGCTGG TCCCTGCTGG CGGTGGGCTC CGCCATTGCC AGTGGCGTGG CCGTGTCCAA GGTGCTGCAC
GCACCTGGGAC AGGTCTTCT CTTTCGGGC GAAGGACCCG AAGGACGACC CGCA^CCGAG CGGATAACGG TCACCGCACC GGCACAGGTT CCACGACGTT
V T L S K K R K R R F L G F L L G V G S A I A S G V A V S K V L H>
a a a a a a a a a a

Figure 3b(i)

26/40

510	CTGGAGGCG	AGGTGAACAA	GATCAAGAGT	GCCCTGCTGT	CCACTAACAA	GGCCGTGGTG	AGCCTGAGCA	ACGGCGTGAG	TGTGCTGACT	AGCAAGGTGC	600
	GACCTCCCGC	TCCACTTCTT	CTAGTTCTCA	CGGGACGACA	GGTGATTGTT	CCGGCACCAC	TCCGACTCGT	TGCCGCATC	ACAGACTGA	TCGTTCCACG	
	L E G	E V N K	I K S	A L L	S T N K	A V V	S L S	N G V	V L T	S K V	
520	a	a	a	a	a	a	a	a	a	a	
530											
540											
550											
560											
570											
580											
590											
600											
610	TGGACCTGAA	GAACTACATC	GACAAGCAAT	TGCTGCCCCAT	CGTGAACAG	CAGTCTGTGTA	GCATCTCCAA	CATCGAGACT	GTGATCGAGT	TCCAGCAGAA	700
	ACCTGGACTT	CTTGATGTAG	CTGTTCTGTTA	ACGACGGGTA	GCACCTGTTT	GTCAGGACAT	CGTAGAGGTT	GTAGCTCTGA	CACTAGCTCA	AGGTCGTCIT	
	L D L	K N Y I	D K Q	L L P I	V N K Q	S C S	I S N I	E T V	I E F	Q Q K	
	a	a	a	a	a	a	a	a	a	a	
620											
630											
640											
650											
660											
670											
680											
690											
700											
710	GAACAACCGC	CTGCTGGAAA	TCACCCGGGA	GTTCAAGTGTG	AACGCTGGCG	TCACCACTCC	TGCTCTCCACC	TACATGCTGA	CCAACAGCGA	GCTGCTGAGC	800
	CTTGTGGCG	GACGACCTTT	AGTGGGCCCT	CAAGTCACAC	TTGGGACCGC	ACTGGTGAGG	ACAGAGGTGG	ATGTACGACT	GGTTGTGCGT	CGACGACTCG	
	N N R	L L E	I T R	E F S	V N A	G V T	T P V	S T Y	M L T	N S E	L L S
	a	a	a	a	a	a	a	a	a	a	
720											
730											
740											
750											
760											
770											
780											
790											
800											
810	CTGATCAACG	ACATGCCCAT	CACCAACGAC	CAGAAGAAGC	TTATGTCCAA	CAACGTGCAG	ATCGTGAGGC	AGCAGAGCTA	CTCCATCATG	AGCATCATCA	900
	GACTAGTTGC	TGTACGGGTA	GTGGTTGCTG	GTCTTCTTCG	AATACAGGTT	GTTGCACGTC	TAGCACTCCG	TCGTCTCGAT	GAGGTAGTAC	TCGTAGTAGT	
	L I N	D M P	I T N	D Q K	K L M	S N V	Q I V	R Q S	Y S I	M S I	I I
	a	a	a	a	a	a	a	a	a	a	
820											
830											
840											
850											
860											
870											
880											
890											
900											
910	AGGAGGAGGT	GCTGGCCTAT	GTGGTGCAGC	TGGCCCTGTA	CGGGGTATC	GATACCCCTT	GCTGGAAGCT	GCACACCAGC	CCCCTGTGCA	CCACCAACAC	1000
	TCCTCCTCCA	CGACCGGATA	CACCACGTGC	ACGGGGACAT	GCCGCAGTAG	CTATGGGAA	CGACCTTCCA	CGTGTGGTGC	GGGACACGT	GGTGGTTGTG	
	K E E	V L A	Y V Q	L P L	Y G V	I D T	P C W	K L H	T S P	L C T	T N T
	a	a	a	a	a	a	a	a	a	a	
920											
930											
940											
950											
960											
970											
980											
990											
1000											

Figure 3b(ii)

27/40

>RsrII

1010 1020 1030 1040 1050 1060 1070 1080 1090 1100
CAAGGAGGC AGCAACATCT GCCTGACCG GACCGACCG GGCTGGTACT GTGACAACGC TGGCTCGGTG AGCTTCTTCC CTCAAGCTGA AACCTGCAAG
GTTCTCTCCG TCGTTGTAGA CCGACTGGC CTGGCTGGC CCGACCATGA CACTGTGG CAGGAGCCAC TCGAAGAAGG GAGTTGACT TTGGACGTTT
K E G S N I C L T R T D R G W Y C D N A G S V S F F P Q A E T C K >

>EcoRV

1110 1120 1130 1140 1150 1160 1170 1180 1190 1200
GTCCAGAGCA ACAGAGTGT CTGTGACACC ATGAACCTCC TGACCTGCC CTCGAGGTG AACCTGTGCA ACGTGGATAT CTTCAACCCC AAGTATGACT
CAGGTCTCGT TGCTCACAA GACACTGTG TACTTGAGG ACTGGGACG GAGGTCCAC TTGGACAGT TGCACCTATA GAAGTTGGG TTCTATCTGA
V Q S N R V F C D T M N S L T L P S E V N L C N V D I F N P K Y D >

>XhoI

1210 1220 1230 1240 1250 1260 1270 1280 1290 1300
GCAAGATCAT GACCTCCAAG ACCGATGTCT CGAGCTCCGT GATCACCAGC CTGGGCGCCA TCGTGTCTG CTATGGCAAG ACCAAGTGCA CCGCCAGCAA
CGTTCTAGTA CTGGAGGTTT TGGCTACAGA GCTCGAGGCA CTAGTGGTGC GACCCGCGT AGCAGAGAC GATACCGTTC TGTTTACGT GCGGTGCTT
C K I M T S K T D V S S S V I T S L G A I V S C Y G K T K C T A S N >

>BstBI

1310 1320 1330 1340 1350 1360 1370 1380 1390 1400
CAAGRACCGG GGCATCATCA AGACCTTCAG CAATGGGTGC GACTACGTTT CGAACAAAGG CGTGGACACT GTGTCCGTGG GCAACACCT GTACTACGTC
GTTCTTGGCC CCGTAGTAGT TCTGGAAGTC GTTACCCAGC CTGATGCAA GCTTGTCCC GCACCTGTGA CACAGGACC CGTTGTGGG CATGATGCAC
K N R G I I K T F S N G C D Y V S N K G V D T V S V G N T L Y Y >

Figure 3b(iii)

```
>XbaI      |      |      >BamHI  
            |      |  
            |      | 1610  
GTCTAGAGGA TCGTCTAGAGGA TC  
CAGATCTCCT AC
```

Figure 3b(iv)

Assembly of F synthetic DNA fragments

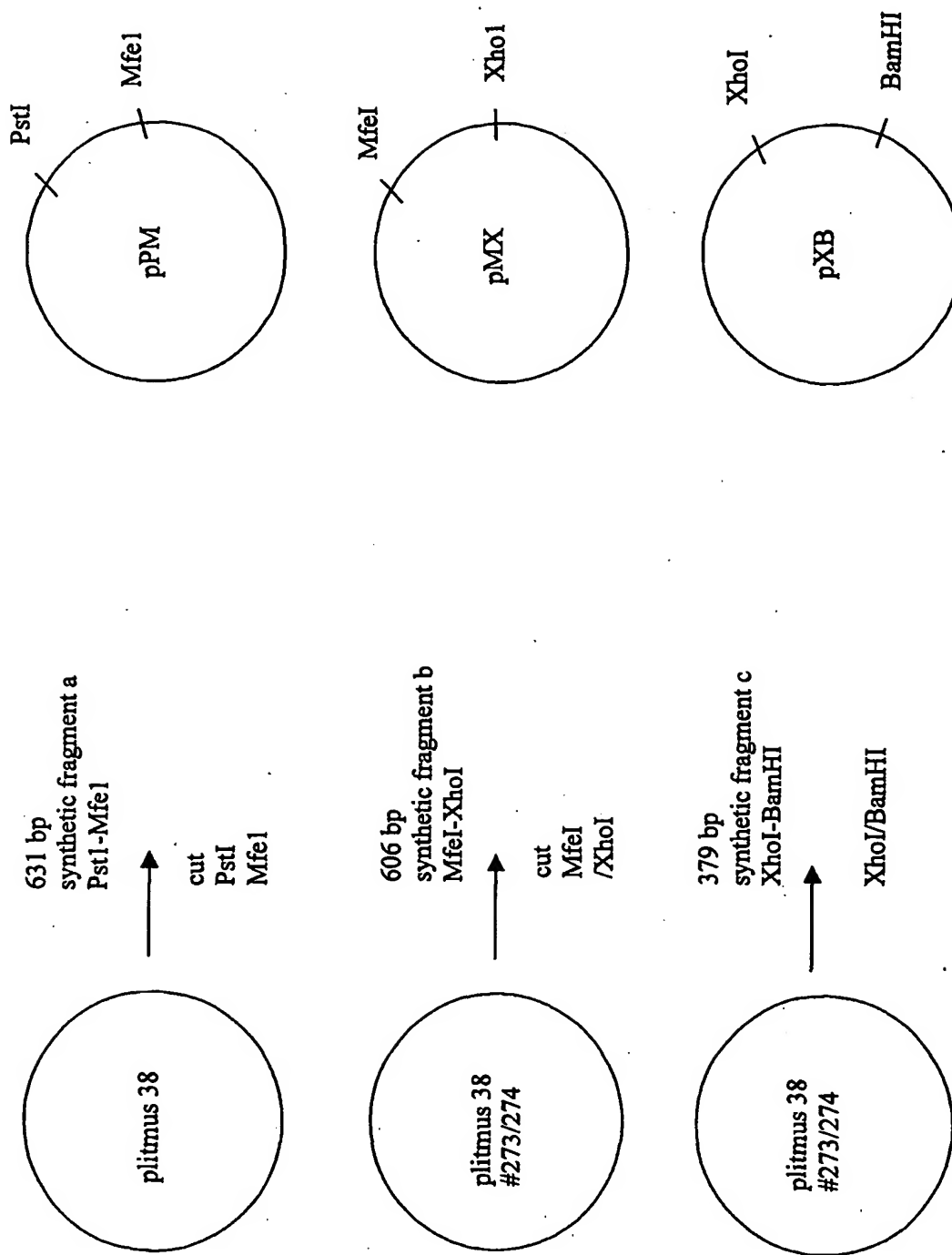


Figure 4a

30/40

Assembly of F in pCICO expression vector

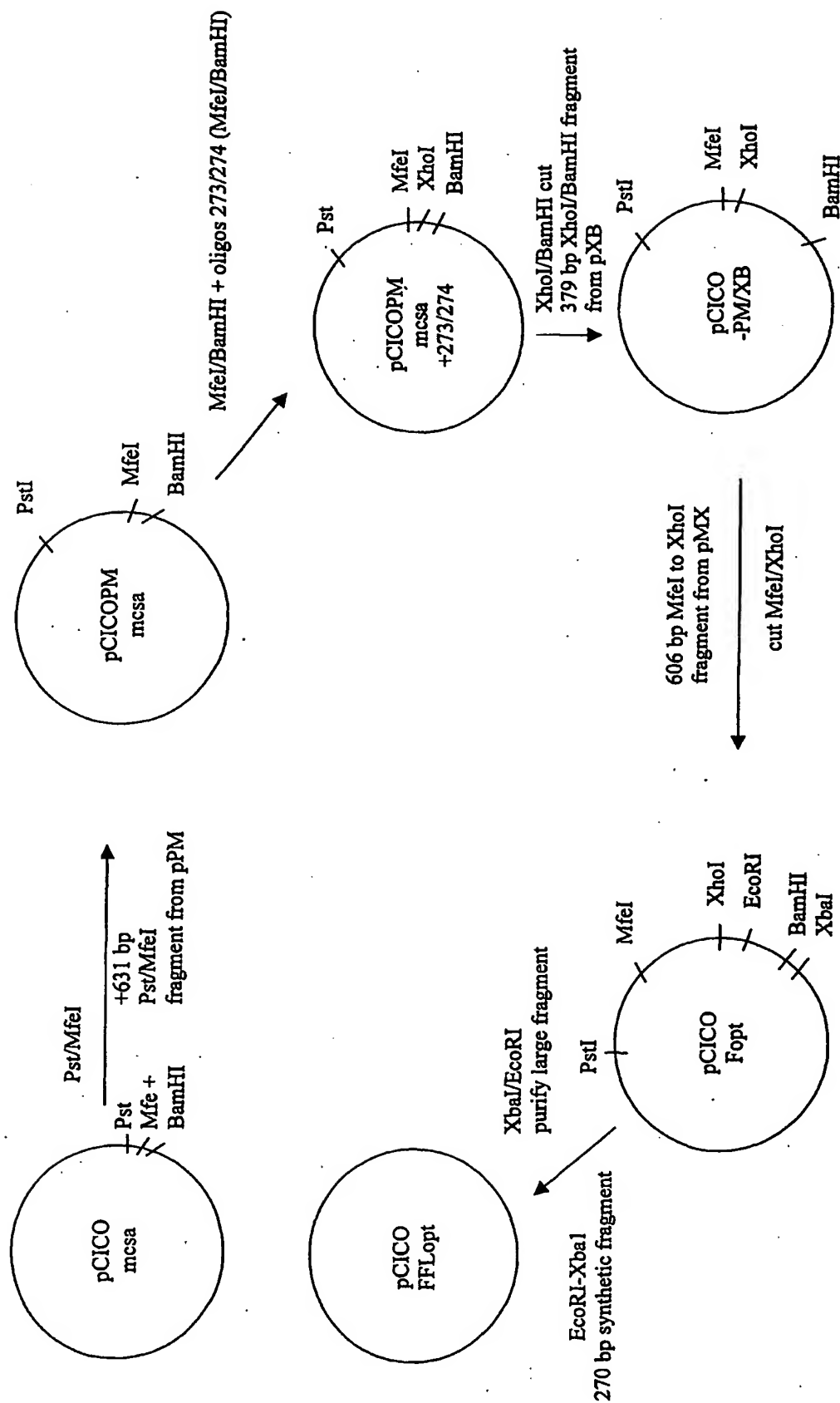


Figure 4b

31/40

**Autoradiograph of immunoprecipitation of
35-S labelled transfected cell supernatants**

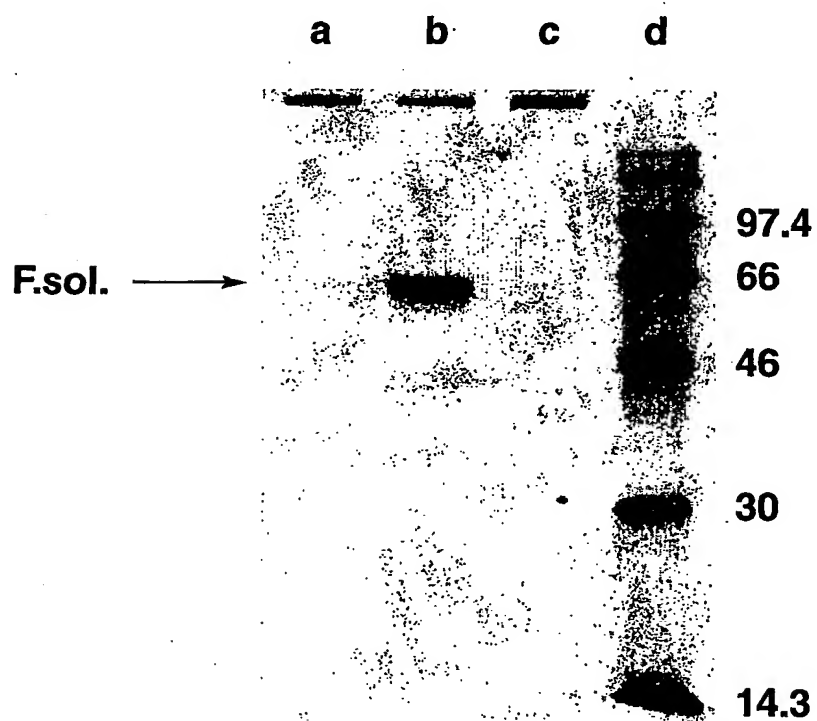


Figure 5

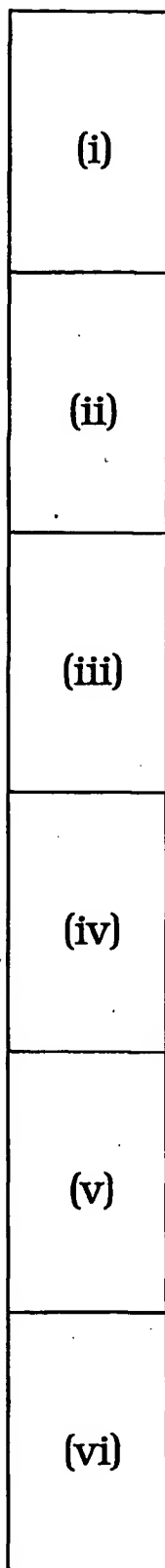


Figure 6
SUBSTITUTE SHEET (RULE 26)

33/40

ClustalW Formatted Alignments

		10	20	30
<i>F.viral</i>		A T G G A G T T G C T A A T C C T C A A A G C A A A T G C A		
<i>F.nat</i>		A T G G A G T T G C T A A T C C T C A A A G C A A A T G C A		
		A T G G A G T T G C T A A T C C T C A A A G C A A A T G C A		
		40	50	60
<i>F.viral</i>		A T T A C C A C A A T C G T C A C T G C A G T C A C A T T T		
<i>F.nat</i>		A T T A C C A C A A T C G T C A C T G C A G T C A C A T T T		
		A T T A C C A C A A T C G T C A C T G C A G T C A C A T T T		
		70	80	90
<i>F.viral</i>		T G T T T T G C T T C T G G T C A A A A C A T C A C T G A A		
<i>F.nat</i>		T G T T T T G C T T C T G G T C A A A A C A T C A C T G A A		
		T G T T T T G C T T C T G G T C A A A A C A T C A C T G A A		
		100	110	120
<i>F.viral</i>		G A A T T T T A T C A A T C A A C A T G C A G T G C A G T T		
<i>F.nat</i>		G A A T T T T A T C A A T C A A C A T G C A G T G C A G T T		
		G A A T T T T A T C A A T C A A C A T G C A G T G C A G T T		
		130	140	150
<i>F.viral</i>		A G C A A A G G C T A T C T T A G T G C T C T G A G A A C T		
<i>F.nat</i>		A G C A A A G G C T A T C T T A G T G C T C T G A G A A C T		
		A G C A A A G G C T A T C T T A G T G C T C T G A G A A C T		
		160	170	180
<i>F.viral</i>		G G T T G G T A T A C C A G T G T T A T A A C T A T A G A A		
<i>F.nat</i>		G G T T G G T A T A C C A G T G T T A T A A C C A T A G A A		
		G G T T G G T A T A C C A G T G T T A T A A C A T A G A A		
		190	200	210
<i>F.viral</i>		T T A A G T A A T A T C A A G A A A A A T A A G T G T A A T		
<i>F.nat</i>		C T A A G T A A T A T C A A G A A A A A T A A G T G T A A T		
		T A A G T A A T A T C A A G A A A A A T A A G T G T A A T		
		220	230	240
<i>F.viral</i>		G G A A C A G A T G C T A A G G T A A A A T T G A T A A A A		
<i>F.nat</i>		G G A A C A G A T G C C A A G G T A A A A T T G A T A A A A		
		G G A A C A G A T G C A A G G T A A A A T T G A T A A A A		
		250	260	270
<i>F.viral</i>		C A A G A A T T A G A T A A A T A T A A A A A T G C T G T A		
<i>F.nat</i>		C A A G A A T T A G A T A A A T A T A A A A A T G C T G T A		
		C A A G A A T T A G A T A A A T A T A A A A A T G C T G T A		
		280	290	300
<i>F.viral</i>		A C A G A A T T G C A G T T G C T C A T G C A A A G C A C A		
<i>F.nat</i>		A C A G A A T T G C A G T T G C T C A T G C A A A G C A C A		
		A C A G A A T T G C A G T T G C T C A T G C A A A G C A C A		

Figure 6(i)

SUBSTITUTE SHEET (RULE 26)

34/40

	310	320	330
<i>F.viral</i>	C A A G C A A C A A A C A A T C G A G C C A G A A G A G A A		
<i>F.nat</i>	C A A G C A A C A A A C A A T C G A G C C A G A A G A G A A		
	340	350	360
<i>F.viral</i>	C T A C C A A G G T T T A T G A A T T A T A C A C T C A A C		
<i>F.nat</i>	C T A C C A A G G T T T A T G A A T T A T A C A C T C A A C		
	370	380	390
<i>F.viral</i>	A A T G C C A A A A A A A C C A A T G T A A C A T T A A G C		
<i>F.nat</i>	A A T G C C A A A A A A A C C A A T G T A A C A T T A A G C		
	400	410	420
<i>F.viral</i>	A A G A A A A G G A A A A G A A G A T T T C T T G G T T T T		
<i>F.nat</i>	A A G A A A A G G A A A A G A A G A T T T C T T G G T T T T		
	430	440	450
<i>F.viral</i>	T T G T T A G G T G T T G G A T C T G C A A T C G C C A G T		
<i>F.nat</i>	T T G T T A G G T G T T G G A T C T G C A A T C G C C A G T		
	460	470	480
<i>F.viral</i>	G G C G T T G C T G T A T C T A A G G T C C T G C A C C T A		
<i>F.nat</i>	G G C G T T G C T G T A T C T A A G G T C C T G C A C C T A		
	490	500	510
<i>F.viral</i>	G A A G G G G A A G T G A A C A A G A T C A A A A G T G C T		
<i>F.nat</i>	G A A G G G G A A G T G A A C A A G A T C A A A A G T G C T		
	520	530	540
<i>F.viral</i>	C T A C T A T C C A C A A A C A A G G G C T G T A G T C A G C		
<i>F.nat</i>	C T A C T A T C C A C A A A C A A G G G C T G T A G T C A G C		
	550	560	570
<i>F.viral</i>	T T A T C A A A T G G A G T T A G T G T C T T A A C C A G C		
<i>F.nat</i>	T T A T C A A A T G G A G T T A G T G T C T T A A C C A G C		
	580	590	600
<i>F.viral</i>	A A A G T G T T A G A C C T C A A A A A C T A T A T A G A T		
<i>F.nat</i>	A A A G T G T T A G A C C T C A A A A A C T A T A T A G A T		
	610	620	630
<i>F.viral</i>	A A A C A A T T G T T A C C T A T T G T G A A C A A G C A A		
<i>F.nat</i>	A A A C A A T T G T T A C C T A T T G T G A A C A A G C A A		

Figure 6(ii)

SUBSTITUTE SHEET (RULE 26)

35/40

	640	650	660
<i>F.viral</i>	A G C T G C A G C A T A T C A A A T A T A G A A A C T G T G		
<i>F.nat</i>	A G C T G C A G C A T A T C A A A T A T A G A A A C T G T G		
	670	680	690
<i>F.viral</i>	A T A G A G T T C C A A C A A A A G A A C A A C A G A C T A		
<i>F.nat</i>	A T A G A G T T C C A A C A A A A G A A C A A C A G A C T A		
	700	710	720
<i>F.viral</i>	C T A G A G A T T A C C A G G G A A T T T A G T G T T A A T		
<i>F.nat</i>	C T A G A G A T T A C C A G G G A A T T T A G T G T T A A T		
	730	740	750
<i>F.viral</i>	G C A G G T G T A A C T A C A C C T G T A A G C A C T T A C		
<i>F.nat</i>	G C A G G T G T A A C T A C A C C T G T A A G C A C T T A C		
	760	770	780
<i>F.viral</i>	A T G T T A A C T A A T A G T G A A T T A T T G T C A T T A		
<i>F.nat</i>	A T G T T A A C T A A T A G T G A A T T A T T G T C A T T A		
	790	800	810
<i>F.viral</i>	A T C A A T G A T A T G C C T A T A A C A A A T G A T C A G		
<i>F.nat</i>	A T C A A T G A T A T G C C T A T A A C A A A T G A T C A G		
	820	830	840
<i>F.viral</i>	A A A A A G T T A A T G T C C A A C A A T G T T C A A A T A		
<i>F.nat</i>	A A A A A G T T A A T G T C C A A C A A T G T T C A A A T A		
	850	860	870
<i>F.viral</i>	G T T A G A C A G C A A A G T T A C T C T A T C A T G T C C		
<i>F.nat</i>	G T T A G A C A G C A A A G T T A C T C T A T C A T G T C C		
	880	890	900
<i>F.viral</i>	A T A A T A A A A G A G G A A G T C T T A G C A T A T G T A		
<i>F.nat</i>	A T A A T A A A A G A G G A A G T C T T A G C A T A T G T A		
	910	920	930
<i>F.viral</i>	G T A C A A T T A C C A C T A T A T G G T G T T A T A G A T		
<i>F.nat</i>	G T A C A A T T A C C A C T A T A T G G T G T T A T A G A T		
	940	950	960
<i>F.viral</i>	A C A C C C T G T T G G A A A C T A C A C A C A T C C C C T		
<i>F.nat</i>	A C A C C C T G T T G G A A A C T A C A C A C A T C C C C T		

Figure 6(iii)

SUBSTITUTE SHEET (RULE 26)

36/40

		970	980	990
<i>F.viral</i>		CTATGTACAACCAACACAAAAGAAGGGTCC		
<i>F.nat</i>		CTATGTACAACCAACACAAAAGAAGGGTCC		
		CTATGTACAACCAACACAAAAGAAGGGTCC		
		1000	1010	1020
<i>F.viral</i>		AACATCTGTTTAAACAAGAACTGACAGAGGA		
<i>F.nat</i>		AACATCTGTTTAAACAAGAACTGACAGAGGA		
		AACATCTGTTTAAACAAGAACTGACAGAGGA		
		1030	1040	1050
<i>F.viral</i>		TGGTACTGTGACAATGCAGGATCAGTATCT		
<i>F.nat</i>		TGGTACTGTGACAATGCAGGATCAGTATCT		
		TGGTACTGTGACAATGCAGGATCAGTATCT		
		1060	1070	1080
<i>F.viral</i>		TTCTTCCCACAAGCTGAAACATGTAAAGTT		
<i>F.nat</i>		TTCTTCCCACAAGCTGAAACATGTAAAGTT		
		TTCTTCCCACAAGCTGAAACATGTAAAGTT		
		1090	1100	1110
<i>F.viral</i>		CAATCAAATCGAGTATTTTGTGACACAATG		
<i>F.nat</i>		CAATCAAATCGAGTATTTTGTGACACAATG		
		CAATCAAATCGAGTATTTTGTGACACAATG		
		1120	1130	1140
<i>F.viral</i>		AACAGTTTAAACATTACCAAGTGAAGTAAAT		
<i>F.nat</i>		AACAGTTTAAACATTACCAAGTGAAGTAAAT		
		AACAGTTTAAACATTACCAAGTGAAGTAAAT		
		1150	1160	1170
<i>F.viral</i>		CTCTGCAATGTTGACATATTCAACCCCAA		
<i>F.nat</i>		CTCTGCAATGTTGACATATTCAACCCCAA		
		CTCTGCAATGTTGACATATTCAACCCCAA		
		1180	1190	1200
<i>F.viral</i>		TATGATTGTAAAATTATGACTTCAAAAACA		
<i>F.nat</i>		TATGATTGTAAAATTATGACTTCAAAAACA		
		TATGATTGTAAAATTATGACTTCAAAAACA		
		1210	1220	1230
<i>F.viral</i>		GATGTAAGCAGCTCCGTTATCACATCTCTA		
<i>F.nat</i>		GATGTAAGCAGCTCCGTTATCACATCTCTA		
		GATGTAAGCAGCTCCGTTATCACATCTCTA		
		1240	1250	1260
<i>F.viral</i>		GGAGCCATTGTGTTCATGCTATGGCAAAACT		
<i>F.nat</i>		GGAGCCATTGTGTTCATGCTATGGCAAAACT		
		GGAGCCATTGTGTTCATGCTATGGCAAAACT		
		1270	1280	1290
<i>F.viral</i>		AAATGTACAGCATCCAATAAAAAATCGTGGA		
<i>F.nat</i>		AAATGTACAGCATCCAATAAAAAATCGTGGA		
		AAATGTACAGCATCCAATAAAAAATCGTGGA		

Figure 6(iv)

SUBSTITUTE SHEET (RULE 26)

37/40

		1300		1310		1320
<i>F.viral</i>		ATCATAAAGACATTTTCTAACGGGTGCGAT				
<i>F.nat</i>		ATCATAAAGACATTTTCTAACGGGTGCGAT				
		ATCATAAAGACATTTTCTAACGGGTGCGAT				
		1330		1340		1350
<i>F.viral</i>		TATGTATCAAATAAAGGGGGTGGACACTGTG				
<i>F.nat</i>		TATGTATCAAATAAAGGGGGTGGACACTGTG				
		TATGTATCAAATAAAGGGGGTGGACACTGTG				
		1360		1370		1380
<i>F.viral</i>		TCTGTAGGTAAACACATTATATTATGTAAAT				
<i>F.nat</i>		TCTGTAGGTAAACACATTATATTATGTAAAT				
		TCTGTAGGTAAACACATTATATTATGTAAAT				
		1390		1400		1410
<i>F.viral</i>		AAGCAAGAAGGTAAAAGTCTCTATGTAAAA				
<i>F.nat</i>		AAGCAAGAAGGTAAAAGTCTCTATGTAAAA				
		AAGCAAGAAGGTAAAAGTCTCTATGTAAAA				
		1420		1430		1440
<i>F.viral</i>		GGTGAACCAATAATAAATTCTATGACCCA				
<i>F.nat</i>		GGTGAACCAATAATAAATTCTATGACCCA				
		GGTGAACCAATAATAAATTCTATGACCCA				
		1450		1460		1470
<i>F.viral</i>		TTAGTATTCCCTCTGATGAATTGTGATGCA				
<i>F.nat</i>		TTAGTATTCCCTCTGATGAATTGTGATGCA				
		TTAGTATTCCCTCTGATGAATTGTGATGCA				
		1480		1490		1500
<i>F.viral</i>		TCAATATCTCAAGTCAACGAGAAAGATTAAAC				
<i>F.nat</i>		TCAATATCTCAAGTCAACGAGAAAGATTAAAC				
		TCAATATCTCAAGTCAACGAGAAAGATTAAAC				
		1510		1520		1530
<i>F.viral</i>		CAGAGCCTAGCATTTATTCGTAAATCCGAT				
<i>F.nat</i>		CAGAGCCTAGCATTTATTCGTAAATCCGAT				
		CAGAGCCTAGCATTTATTCGTAAATCCGAT				
		1540		1550		1560
<i>F.viral</i>		GAATTATTACATAATGTAAATGCTGGTAAA				
<i>F.nat</i>		GAATTATTACATAATGTAAATGCTGGTAAA				
		GAATTATTACATAATGTAAATGCTGGTAAA				
		1570		1580		1590
<i>F.viral</i>		TCCACCACAAATATCATGATAACTACTATA				
<i>F.nat</i>		TCCACCACAAATATCATGATAACTACTATA				
		TCCACCACAAATATCATGATAACTACTATA				
		1600		1610		1620
<i>F.viral</i>		ATTATAGTGATTATAGTAATATTGTTATCA				
<i>F.nat</i>		ATTATAGTGATTATAGTAATATTGTTATCA				
		ATTATAGTGATTATAGTAATATTGTTATCA				

Figure 6(v)

SUBSTITUTE SHEET (RULE 26)

38/40

	1630	1640	1650
<i>F.viral</i>	T T A A T T G C T G T T G G A C T G C T C T T A T A C T G T		
<i>F.nat</i>	T T A A T T G C T G T T G G A C T G C T C T T A T A C T G T		
	1660	1670	1680
<i>F.viral</i>	A A G G C C A G A A G C A C A C C A G T C A C A C T A A G C		
<i>F.nat</i>	A A G G C C A G A A G C A C A C C A G T C A C A C T A A G C		
	1690	1700	1710
<i>F.viral</i>	A A A G A T C A A C T G A G T G G T A T A A A T A A T A T T		
<i>F.nat</i>	A A A G A T C A A C T G A G T G G T A T A A A T A A T A T T		
	1720	1730	1740
<i>F.viral</i>	G C A T T T A G T A A C T A A		
<i>F.nat</i>	G C A T T T A G T A A C T A A		
	G C A T T T A G T A A C T A A		

Figure 6(vi)

SUBSTITUTE SHEET (RULE 26)

39/40

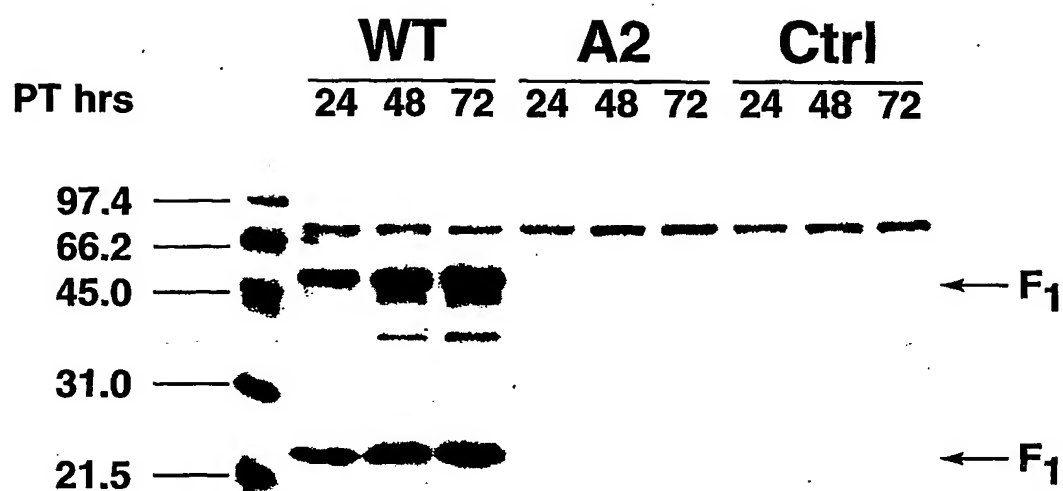


Figure 7
SUBSTITUTE SHEET (RULE 26)

40/40

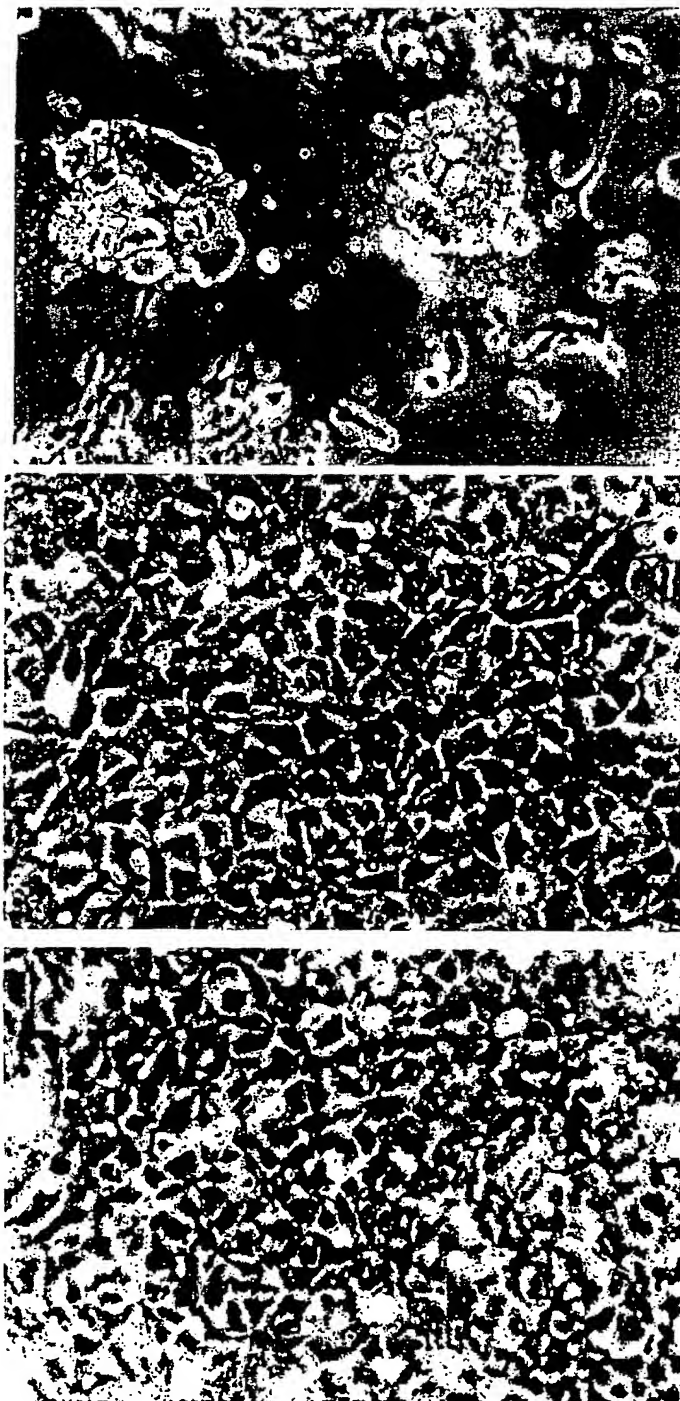


Figure 8
SUBSTITUTE SHEET (RULE 26)